

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
3 January 2003 (03.01.2003)

PCT

(10) International Publication Number  
WO 03/000113 A2

- (51) International Patent Classification: **A61B**
- (21) International Application Number: PCT/US02/19592
- (22) International Filing Date: 19 June 2002 (19.06.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |            |                                |    |
|------------|--------------------------------|----|
| 60/299,500 | 20 June 2001 (20.06.2001)      | US |
| 60/300,889 | 25 June 2001 (25.06.2001)      | US |
| 60/301,280 | 29 June 2001 (29.06.2001)      | US |
| 60/304,813 | 11 July 2001 (11.07.2001)      | US |
| 60/312,312 | 13 August 2001 (13.08.2001)    | US |
| 60/314,280 | 22 August 2001 (22.08.2001)    | US |
| 60/323,268 | 18 September 2001 (18.09.2001) | US |
| 60/330,227 | 19 October 2001 (19.10.2001)   | US |
| 60/336,827 | 7 November 2001 (07.11.2001)   | US |
| 60/366,869 | 20 March 2002 (20.03.2002)     | US |
| 60/378,883 | 8 May 2002 (08.05.2002)        | US |

(71) Applicant: GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors: CAIRNS, Belinda; 1229 Paloma Avenue, Burlingame, CA 94010 (US). CHEN, Ruihuan; #301, 1110 Polynesian Drive, Foster City, CA 94044 (US). FRANTZ, Gretchen; 135 San Benito Way, San Francisco, CA 94127 (US). HILLAN, Kenneth, J.; 64 Seward Street, San Francisco, CA 94114 (US). KOEPPEN, Hartmut; 1746 Milvia Street, Berkeley, CA 94709 (US). PHILLIPS, Heidi, S.; 15 Pine Avenue, San Carlos, CA 94070 (US). POLAKIS, Paul; 1449 Cortez Avenue,

Burlingame, CA 94010 (US). SMITH, Victoria; 19 Dwight Road, Burlingame, CA 94010 (US). SPENSER, Susan, B.; 2495 Paradise Drive, Tiburon, CA 94920 (US). WILLIAMS, P., Mickey; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WU, Thomas, D.; 113 Elsie Street, San Francisco, CA 94116 (US). ZHANG, Zemin; 876 Taurus Drive, Foster City, CA 94044 (US).

(74) Agent: KRESNAK, Mark, T. et al.; Genentech, Inc., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States (national): AT, AG, AL, AM, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

## Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

(57) Abstract: The present invention is directed to compositions of matter useful for the diagnosis and treatment of tumor in mammals and to methods of using those compositions of matter for the same.

WO 03/000113 A2

## COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

FIELD OF THE INVENTION

The present invention is directed to compositions of matter useful for the diagnosis and treatment of tumor in mammals and to methods of using those compositions of matter for the same.

BACKGROUND OF THE INVENTION

Malignant tumors (cancers) are the second leading cause of death in the United States, after heart disease (Boring et al., *CA Cancer J. Clin.* 43:7 (1993)). Cancer is characterized by the increase in the number of abnormal, or neoplastic, cells derived from a normal tissue which proliferate to form a tumor mass, the invasion of adjacent tissues by these neoplastic tumor cells, and the generation of malignant cells which eventually spread via the blood or lymphatic system to regional lymph nodes and to distant sites via a process called metastasis. In a cancerous state, a cell proliferates under conditions in which normal cells would not grow. Cancer manifests itself in a wide variety of forms, characterized by different degrees of invasiveness and aggressiveness.

In attempts to discover effective cellular targets for cancer diagnosis and therapy, researchers have sought to identify transmembrane or otherwise membrane-associated polypeptides that are specifically expressed on the surface of one or more particular type(s) of cancer cell as compared to on one or more normal non-cancerous cell(s). Often, such membrane-associated polypeptides are more abundantly expressed on the surface of the cancer cells as compared to on the surface of the non-cancerous cells. The identification of such tumor-associated cell surface antigen polypeptides has given rise to the ability to specifically target cancer cells for destruction via antibody-based therapies. In this regard, it is noted that antibody-based therapy has proved very effective in the treatment of certain cancers. For example, HERCEPTIN® and RITUXAN® (both from Genentech Inc., South San Francisco, California) are antibodies that have been used successfully to treat breast cancer and non-Hodgkin's lymphoma, respectively. More specifically, HERCEPTIN® is a recombinant DNA-derived humanized monoclonal antibody that selectively binds to the extracellular domain of the human epidermal growth factor receptor 2 (HER2) proto-oncogene. HER2 protein overexpression is observed in 25-30% of primary breast cancers. RITUXAN® is a genetically engineered chimeric murine/human monoclonal antibody directed against the CD20 antigen found on the surface of normal and malignant B lymphocytes. Both these antibodies are recombinantly produced in CHO cells.

In other attempts to discover effective cellular targets for cancer diagnosis and therapy, researchers have sought to identify (1) non-membrane-associated polypeptides that are specifically produced by one or more particular type(s) of cancer cell(s) as compared to by one or more particular type(s) of non-cancerous normal cell(s), (2) polypeptides that are produced by cancer cells at an expression level that is significantly higher than that of one or more normal non-cancerous cell(s), or (3) polypeptides whose expression is

specifically limited to only a single (or very limited number of different) tissue type(s) in both the cancerous and non-cancerous state (e.g., normal prostate and prostate tumor tissue). Such polypeptides may remain intracellularly located or may be secreted by the cancer cell. Moreover, such polypeptides may be expressed not by the cancer cell itself, but rather by cells which produce and/or secrete polypeptides having a potentiating or growth-enhancing effect on cancer cells. Such secreted polypeptides are often proteins that provide cancer cells with a growth advantage over normal cells and include such things as, for example, angiogenic factors, cellular adhesion factors, growth factors, and the like. Identification of antagonists of such non-membrane associated polypeptides would be expected to serve as effective therapeutic agents for the treatment of such cancers. Furthermore, identification of the expression pattern of such polypeptides would be useful for the diagnosis of particular cancers in mammals.

Despite the above identified advances in mammalian cancer therapy, there is a great need for additional diagnostic and therapeutic agents capable of detecting the presence of tumor in a mammal and for effectively inhibiting neoplastic cell growth, respectively. Accordingly, it is an objective of the present invention to identify: (1) cell membrane-associated polypeptides that are more abundantly expressed on one or more type(s) of cancer cell(s) as compared to on normal cells or on other different cancer cells, (2) non-membrane-associated polypeptides that are specifically produced by one or more particular type(s) of cancer cell(s) (or by other cells that produce polypeptides having a potentiating effect on the growth of cancer cells) as compared to by one or more particular type(s) of non-cancerous normal cell(s), (3) non-membrane-associated polypeptides that are produced by cancer cells at an expression level that is significantly higher than that of one or more normal non-cancerous cell(s), or (4) polypeptides whose expression is specifically limited to only a single (or very limited number of different) tissue type(s) in both a cancerous and non-cancerous state (e.g., normal prostate and prostate tumor tissue), and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of cancer in mammals. It is also an objective of the present invention to identify cell membrane-associated, secreted or intracellular polypeptides whose expression is limited to a single or very limited number of tissues, and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of cancer in mammals.

## SUMMARY OF THE INVENTION

### A. Embodiments

In the present specification, Applicants describe for the first time the identification of various cellular polypeptides (and their encoding nucleic acids or fragments thereof) which are expressed to a greater degree on the surface of or by one or more types of cancer cell(s) as compared to on the surface of or by one or more types of normal non-cancer cells. Alternatively, such polypeptides are expressed by cells which produce and/or secrete polypeptides having a potentiating or growth-enhancing effect on cancer cells. Again alternatively, such polypeptides may not be overexpressed by tumor cells as compared to normal cells of the same tissue type, but rather may be specifically expressed by both tumor cells and normal cells of only a single or very limited number of tissue types (preferably tissues which are not essential for life, e.g., prostate, etc.).

All of the above polypeptides are herein referred to as Tumor-associated Antigenic Target polypeptides ("TAT" polypeptides) and are expected to serve as effective targets for cancer therapy and diagnosis in mammals.

Accordingly, in one embodiment of the present invention, the invention provides an isolated nucleic acid molecule having a nucleotide sequence that encodes a tumor-associated antigenic target polypeptide or fragment thereof (a "TAT" polypeptide).

In certain aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule encoding a full-length TAT polypeptide having an amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule comprising the coding sequence of a full-length TAT polypeptide cDNA as disclosed herein, the coding sequence of a TAT polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In further aspects, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule that encodes the same mature polypeptide encoded by the full-length coding region of any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect of the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a TAT polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide(s) are disclosed herein. Therefore, soluble extracellular domains of the herein described TAT polypeptides are contemplated.

In other aspects, the present invention is directed to isolated nucleic acid molecules which hybridize to (a) a nucleotide sequence encoding a TAT polypeptide having a full-length amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein,



or (b) the complement of the nucleotide sequence of (a). In this regard, an embodiment of the present invention is directed to fragments of a full-length TAT polypeptide coding sequence, or the complement thereof, as disclosed herein, that may find use as, for example, hybridization probes useful as, for example, diagnostic probes, antisense oligonucleotide probes, or for encoding fragments of a full-length TAT polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-TAT polypeptide antibody, a TAT binding oligopeptide or other small organic molecule that binds to a TAT polypeptide. Such nucleic acid fragments are usually at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a TAT polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the TAT polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which TAT polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such novel fragments of TAT polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the TAT polypeptide fragments encoded by these nucleotide molecule fragments, preferably those TAT polypeptide fragments that comprise a binding site for an anti-TAT antibody, a TAT binding oligopeptide or other small organic molecule that binds to a TAT polypeptide.

In another embodiment, the invention provides isolated TAT polypeptides encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated TAT polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity, to a TAT polypeptide having a full-length amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide protein, with or without the signal peptide, as disclosed herein, an amino acid sequence encoded by any of the nucleic acid sequences disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated TAT polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated TAT polypeptide without the N-terminal signal sequence and/or without the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the TAT polypeptide and recovering the TAT polypeptide from the cell culture.

Another aspect of the invention provides an isolated TAT polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the TAT polypeptide and recovering the TAT polypeptide from the cell culture.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cells comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli* cells, or yeast cells. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides isolated chimeric polypeptides comprising any of the herein described TAT polypeptides fused to a heterologous (non-TAT) polypeptide. Example of such chimeric molecules comprise any of the herein described TAT polypeptides fused to a heterologous polypeptide such as, for example, an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, single-chain antibody or antibody that competitively inhibits the binding of an anti-TAT polypeptide antibody to its respective antigenic epitope. Antibodies of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies of the present invention may optionally be produced in CHO cells or bacterial cells and preferably induce death of a cell to which they bind. For diagnostic purposes, the antibodies of the present invention may be detectably labeled, attached to a solid support, or the like.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described antibodies. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli* cells, or yeast cells. A process for producing any of the herein described antibodies is further provided and comprises culturing host cells under conditions suitable for expression of the desired antibody and recovering the desired antibody from the cell culture.

In another embodiment, the invention provides oligopeptides ("TAT binding oligopeptides") which bind, preferably specifically, to any of the above or below described TAT polypeptides. Optionally, the TAT binding oligopeptides of the present invention may be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive

isotope, a nucleolytic enzyme, or the like. The TAT binding oligopeptides of the present invention may optionally be produced in CHO cells or bacterial cells and preferably induce death of a cell to which they bind. For diagnostic purposes, the TAT binding oligopeptides of the present invention may be detectably labeled, attached to a solid support, or the like.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described TAT binding oligopeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli* cells, or yeast cells. A process for producing any of the herein described TAT binding oligopeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired oligopeptide and recovering the desired oligopeptide from the cell culture.

In another embodiment, the invention provides small organic molecules ("TAT binding organic molecules") which bind, preferably specifically, to any of the above or below described TAT polypeptides. Optionally, the TAT binding organic molecules of the present invention may be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The TAT binding organic molecules of the present invention preferably induce death of a cell to which they bind. For diagnostic purposes, the TAT binding organic molecules of the present invention may be detectably labeled, attached to a solid support, or the like.

In a still further embodiment, the invention concerns a composition of matter comprising a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

In yet another embodiment, the invention concerns an article of manufacture comprising a container and a composition of matter contained within the container, wherein the composition of matter may comprise a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein. The article may further optionally comprise a label affixed to the container, or a package insert included with the container, that refers to the use of the composition of matter for the therapeutic treatment or diagnostic detection of a tumor.

Another embodiment of the present invention is directed to the use of a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT polypeptide antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein, for the preparation of a medicament useful in the treatment of a condition which is responsive to the TAT polypeptide, chimeric TAT polypeptide, anti-TAT polypeptide antibody, TAT binding oligopeptide, or TAT binding organic molecule.

## B. Additional Embodiments

Another embodiment of the present invention is directed to a method for inhibiting the growth of a cell that expresses a TAT polypeptide, wherein the method comprises contacting the cell with an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, and wherein the binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide causes inhibition of the growth of the cell expressing the TAT polypeptide. In preferred embodiments, the cell is a cancer cell and binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide causes death of the cell expressing the TAT polypeptide. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and TAT binding oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of therapeutically treating a mammal having a cancerous tumor comprising cells that express a TAT polypeptide, wherein the method comprises administering to the mammal a therapeutically effective amount of an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby resulting in the effective therapeutic treatment of the tumor. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of determining the presence of a TAT polypeptide in a sample suspected of containing the TAT polypeptide, wherein the method comprises exposing the sample to an antibody, oligopeptide or small organic molecule that binds to the TAT polypeptide and determining binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide in the sample, wherein the presence of such binding is indicative of the presence of the TAT polypeptide in the sample. Optionally, the sample may contain cells (which may be cancer cells) suspected of expressing the TAT polypeptide. The antibody, TAT binding oligopeptide or TAT binding organic molecule employed in the method may optionally be detectably labeled, attached to a solid support, or the like.

A further embodiment of the present invention is directed to a method of diagnosing the presence of a tumor in a mammal, wherein the method comprises detecting the level of expression of a gene encoding a TAT polypeptide (a) in a test sample of tissue cells obtained from said mammal, and (b) in a control sample of known normal non-cancerous cells of the same tissue origin or type, wherein a higher level of expression of the TAT polypeptide in the test sample, as compared to the control sample, is indicative of the presence of

tumor in the mammal from which the test sample was obtained.

Another embodiment of the present invention is directed to a method of diagnosing the presence of a tumor in a mammal, wherein the method comprises (a) contacting a test sample comprising tissue cells obtained from the mammal with an antibody, oligopeptide or small organic molecule that binds to a TAT polypeptide and (b) detecting the formation of a complex between the antibody, oligopeptide or small organic molecule and the TAT polypeptide in the test sample, wherein the formation of a complex is indicative of the presence of a tumor in the mammal. Optionally, the antibody, TAT binding oligopeptide or TAT binding organic molecule employed is detectably labeled, attached to a solid support, or the like, and/or the test sample of tissue cells is obtained from an individual suspected of having a cancerous tumor.

Yet another embodiment of the present invention is directed to a method for treating or preventing a cell proliferative disorder associated with altered, preferably increased, expression or activity of a TAT polypeptide, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT polypeptide. Preferably, the cell proliferative disorder is cancer and the antagonist of the TAT polypeptide is an anti-TAT polypeptide antibody, TAT binding oligopeptide, TAT binding organic molecule or antisense oligonucleotide. Effective treatment or prevention of the cell proliferative disorder may be a result of direct killing or growth inhibition of cells that express a TAT polypeptide or by antagonizing the cell growth potentiating activity of a TAT polypeptide.

Yet another embodiment of the present invention is directed to a method of binding an antibody, oligopeptide or small organic molecule to a cell that expresses a TAT polypeptide, wherein the method comprises contacting a cell that expresses a TAT polypeptide with said antibody, oligopeptide or small organic molecule under conditions which are suitable for binding of the antibody, oligopeptide or small organic molecule to said TAT polypeptide and allowing binding therebetween.

Other embodiments of the present invention are directed to the use of (a) a TAT polypeptide, (b) a nucleic acid encoding a TAT polypeptide or a vector or host cell comprising that nucleic acid, (c) an anti-TAT polypeptide antibody, (d) a TAT-binding oligopeptide, or (e) a TAT-binding small organic molecule in the preparation of a medicament useful for (i) the therapeutic treatment or diagnostic detection of a cancer or tumor, or (ii) the therapeutic treatment or prevention of a cell proliferative disorder.

Another embodiment of the present invention is directed to a method for inhibiting the growth of a cancer cell, wherein the growth of said cancer cell is at least in part dependent upon the growth potentiating effect(s) of a TAT polypeptide (wherein the TAT polypeptide may be expressed either by the cancer cell itself or a cell that produces polypeptide(s) that have a growth potentiating effect on cancer cells), wherein the method comprises contacting the TAT polypeptide with an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby antagonizing the growth-potentiating activity of the TAT polypeptide and, in turn, inhibiting the growth of the cancer cell. Preferably the growth of the cancer cell is completely inhibited. Even more preferably, binding of the antibody, oligopeptide or small organic molecule to the TAT polypeptide induces the death of the cancer cell. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present

invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and TAT binding oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of therapeutically treating a tumor in a mammal, wherein the growth of said tumor is at least in part dependent upon the growth potentiating effect(s) of a TAT polypeptide, wherein the method comprises administering to the mammal a therapeutically effective amount of an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby antagonizing the growth potentiating activity of said TAT polypeptide and resulting in the effective therapeutic treatment of the tumor. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

#### C. Further Additional Embodiments

In yet further embodiments, the invention is directed to the following set of potential claims for this application:

1. Isolated nucleic acid having a nucleotide sequence that has at least 80% nucleic acid sequence identity to:

(a) a DNA molecule encoding the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a DNA molecule encoding the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a DNA molecule encoding an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a DNA molecule encoding an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

2. Isolated nucleic acid having:

(a) a nucleotide sequence that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a nucleotide sequence that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a nucleotide sequence that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

5 (d) a nucleotide sequence that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

10 3. Isolated nucleic acid that hybridizes to:

(a) a nucleic acid that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a nucleic acid that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

15 (c) a nucleic acid that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a nucleic acid that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

20 (f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

4. The nucleic acid of Claim 3, wherein the hybridization occurs under stringent conditions.

5. The nucleic acid of Claim 3 which is at least about 5 nucleotides in length.

25 6. An expression vector comprising the nucleic acid of Claim 1, 2 or 3.

7. The expression vector of Claim 6, wherein said nucleic acid is operably linked to control sequences recognized by a host cell transformed with the vector.

8. A host cell comprising the expression vector of Claim 7.

9. The host cell of Claim 8 which is a CHO cell, an *E. coli* cell or a yeast cell.

30 10. A process for producing a polypeptide comprising culturing the host cell of Claim 8 under conditions suitable for expression of said polypeptide and recovering said polypeptide from the cell culture.

11. An isolated polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

35 (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

5 (f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

12. An isolated polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

10 (c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

15 (e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

13. A chimeric polypeptide comprising the polypeptide of Claim 11 or 12 fused to a heterologous polypeptide.

20 14. The chimeric polypeptide of Claim 13, wherein said heterologous polypeptide is an epitope tag sequence or an Fc region of an immunoglobulin.

15. An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

25 (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

30 (d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

35 16. An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);



(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

17. The antibody of Claim 15 or 16 which is a monoclonal antibody.

18. The antibody of Claim 15 or 16 which is an antibody fragment.

19. The antibody of Claim 15 or 16 which is a chimeric or a humanized antibody.

20. The antibody of Claim 15 or 16 which is conjugated to a growth inhibitory agent.

21. The antibody of Claim 15 or 16 which is conjugated to a cytotoxic agent.

22. The antibody of Claim 21, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

23. The antibody of Claim 21, wherein the cytotoxic agent is a toxin.

24. The antibody of Claim 23, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

25. The antibody of Claim 23, wherein the toxin is a maytansinoid.

26. The antibody of Claim 15 or 16 which is produced in bacteria.

27. The antibody of Claim 15 or 16 which is produced in CHO cells.

28. The antibody of Claim 15 or 16 which induces death of a cell to which it binds.

29. The antibody of Claim 15 or 16 which is detectably labeled.

30. An isolated nucleic acid having a nucleotide sequence that encodes the antibody of Claim 15 or 16.

31. An expression vector comprising the nucleic acid of Claim 30 operably linked to control sequences recognized by a host cell transformed with the vector.

32. A host cell comprising the expression vector of Claim 31.

33. The host cell of Claim 32 which is a CHO cell, an *E. coli* cell or a yeast cell.

34. A process for producing an antibody comprising culturing the host cell of Claim 32 under conditions suitable for expression of said antibody and recovering said antibody from the cell culture.

35. An isolated oligopeptide that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

5 (e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

36. An isolated oligopeptide that binds to a polypeptide having:

10 (a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

15 (d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

20 37. The oligopeptide of Claim 35 or 36 which is conjugated to a growth inhibitory agent.

38. The oligopeptide of Claim 35 or 36 which is conjugated to a cytotoxic agent.

39. The oligopeptide of Claim 38, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

40. The oligopeptide of Claim 38, wherein the cytotoxic agent is a toxin.

25 41. The oligopeptide of Claim 40, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

42. The oligopeptide of Claim 40, wherein the toxin is a maytansinoid.

43. The oligopeptide of Claim 35 or 36 which induces death of a cell to which it binds.

44. The oligopeptide of Claim 35 or 36 which is detectably labeled.

30 45. A TAT binding organic molecule that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

35 (c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

5 (f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

46. The organic molecule of Claim 45 that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

10 (c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

15 (e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

47. The organic molecule of Claim 45 or 46 which is conjugated to a growth inhibitory agent.

48. The organic molecule of Claim 45 or 46 which is conjugated to a cytotoxic agent.

20 49. The organic molecule of Claim 48, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

50. The organic molecule of Claim 48, wherein the cytotoxic agent is a toxin.

51. The organic molecule of Claim 50, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

25 52. The organic molecule of Claim 50, wherein the toxin is a maytansinoid.

53. The organic molecule of Claim 45 or 46 which induces death of a cell to which it binds.

54. The organic molecule of Claim 45 or 46 which is detectably labeled.

55. A composition of matter comprising:

(a) the polypeptide of Claim 11;

30 (b) the polypeptide of Claim 12;

(c) the chimeric polypeptide of Claim 13;

(d) the antibody of Claim 15;

(e) the antibody of Claim 16;

(f) the oligopeptide of Claim 35;

35 (g) the oligopeptide of Claim 36;

(h) the TAT binding organic molecule of Claim 45; or

(i) the TAT binding organic molecule of Claim 46; in combination with a carrier.

56. The composition of matter of Claim 55, wherein said carrier is a pharmaceutically acceptable carrier.
57. An article of manufacture comprising:
- (a) a container; and
  - (b) the composition of matter of Claim 55 contained within said container.
58. The article of manufacture of Claim 57 further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.
59. A method of inhibiting the growth of a cell that expresses a protein having at least 80% amino acid sequence identity to:
- (a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);
  - (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;
  - (c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;
  - (d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;
  - (e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or
  - (f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said cell with an antibody, oligopeptide or organic molecule that binds to said protein, the binding of said antibody, oligopeptide or organic molecule to said protein thereby causing an inhibition of growth of said cell.
60. The method of Claim 59, wherein said antibody is a monoclonal antibody.
61. The method of Claim 59, wherein said antibody is an antibody fragment.
62. The method of Claim 59, wherein said antibody is a chimeric or a humanized antibody.
63. The method of Claim 59, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.
64. The method of Claim 59, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.
65. The method of Claim 64, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.
66. The method of Claim 64, wherein the cytotoxic agent is a toxin.
67. The method of Claim 66, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.
68. The method of Claim 66, wherein the toxin is a maytansinoid.
69. The method of Claim 59, wherein said antibody is produced in bacteria.
70. The method of Claim 59, wherein said antibody is produced in CHO cells.

71. The method of Claim 59, wherein said cell is a cancer cell.

72. The method of Claim 71, wherein said cancer cell is further exposed to radiation treatment or a chemotherapeutic agent.

73. The method of Claim 71, wherein said cancer cell is selected from the group consisting of a breast cancer cell, a colorectal cancer cell, a lung cancer cell, an ovarian cancer cell, a central nervous system cancer cell, a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a cervical cancer cell, a melanoma cell and a leukemia cell.

74. The method of Claim 71, wherein said protein is more abundantly expressed by said cancer cell as compared to a normal cell of the same tissue origin.

75. The method of Claim 59 which causes the death of said cell.

76. The method of Claim 59, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

77. A method of therapeutically treating a mammal having a cancerous tumor comprising cells that express a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising administering to said mammal a therapeutically effective amount of an antibody, oligopeptide or organic molecule that binds to said protein, thereby effectively treating said mammal.

78. The method of Claim 77, wherein said antibody is a monoclonal antibody.

79. The method of Claim 77, wherein said antibody is an antibody fragment.
80. The method of Claim 77, wherein said antibody is a chimeric or a humanized antibody.
81. The method of Claim 77, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.
- 5 82. The method of Claim 77, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.
83. The method of Claim 82, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.
84. The method of Claim 82, wherein the cytotoxic agent is a toxin.
85. The method of Claim 84, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.
- 10 86. The method of Claim 84, wherein the toxin is a maytansinoid.
87. The method of Claim 77, wherein said antibody is produced in bacteria.
88. The method of Claim 77, wherein said antibody is produced in CHO cells.
89. The method of Claim 77, wherein said tumor is further exposed to radiation treatment or a  
15 chemotherapeutic agent.
90. The method of Claim 77, wherein said tumor is a breast tumor, a colorectal tumor, a lung tumor, an ovarian tumor, a central nervous system tumor, a liver tumor, a bladder tumor, a pancreatic tumor, or a cervical tumor.
91. The method of Claim 77, wherein said protein is more abundantly expressed by the  
20 cancerous cells of said tumor as compared to a normal cell of the same tissue origin.
92. The method of Claim 77, wherein said protein has:
- (a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);
- (b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;
- 25 (c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;
- (d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;
- (e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-  
30 B (SEQ ID NOS:1-78); or
- (f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).
93. A method of determining the presence of a protein in a sample suspected of containing said  
protein, wherein said protein has at least 80% amino acid sequence identity to:
- 35 (a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);
- (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

5 (e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising exposing said sample to an antibody, oligopeptide or organic molecule that binds to said protein and determining binding of said antibody, oligopeptide or organic molecule to said protein in said sample, wherein binding of the antibody, oligopeptide or organic molecule to said protein is indicative of the presence of said protein in said sample.

94. The method of Claim 93, wherein said sample comprises a cell suspected of expressing said protein.

95. The method of Claim 94, wherein said cell is a cancer cell.

96. The method of Claim 93, wherein said antibody, oligopeptide or organic molecule is  
15 detectably labeled.

97. The method of Claim 93, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

20 (c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

25 (e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

98. A method of diagnosing the presence of a tumor in a mammal, said method comprising determining the level of expression of a gene encoding a protein having at least 80% amino acid sequence  
30 identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

35 (c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), in a test sample of tissue cells obtained from said mammal and in a control sample of known normal cells of the same tissue origin, wherein a higher level of expression of said protein in the test sample, as compared to the control sample, is indicative of the presence of tumor in the mammal from which the test sample was obtained.

99. The method of Claim 98, wherein the step of determining the level of expression of a gene encoding said protein comprises employing an oligonucleotide in an *in situ* hybridization or RT-PCR analysis.

100. The method of Claim 98, wherein the step of determining the level of expression of a gene encoding said protein comprises employing an antibody in an immunohistochemistry or Western blot analysis.

101. The method of Claim 98, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

102. A method of diagnosing the presence of a tumor in a mammal, said method comprising contacting a test sample of tissue cells obtained from said mammal with an antibody, oligopeptide or organic molecule that binds to a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), and detecting the formation of a complex between said antibody, oligopeptide or organic molecule and said protein in the test sample, wherein the formation of a



complex is indicative of the presence of a tumor in said mammal.

103. The method of Claim 102, wherein said antibody, oligopeptide or organic molecule is detectably labeled.

104. The method of Claim 102, wherein said test sample of tissue cells is obtained from an individual suspected of having a cancerous tumor.

5 105. The method of Claim 102, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

10 (c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

15 (f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

106. A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

20 (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

25 (d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

30 (f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of said protein, thereby effectively treating or preventing said cell proliferative disorder.

107. The method of Claim 106, wherein said cell proliferative disorder is cancer.

108. The method of Claim 106, wherein said antagonist is an anti-TAT polypeptide antibody, TAT binding oligopeptide, TAT binding organic molecule or antisense oligonucleotide.

35 109. A method of binding an antibody, oligopeptide or organic molecule to a cell that expresses a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

5 (d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

10 (f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said cell with an antibody, oligopeptide or organic molecule that binds to said protein and allowing the binding of the antibody, oligopeptide or organic molecule to said protein to occur, thereby binding said antibody, oligopeptide or organic molecule to said cell.

110. The method of Claim 109, wherein said antibody is a monoclonal antibody.

111. The method of Claim 109, wherein said antibody is an antibody fragment.

15 112. The method of Claim 109, wherein said antibody is a chimeric or a humanized antibody.

113. The method of Claim 109, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

114. The method of Claim 109, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

20 115. The method of Claim 114, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

116. The method of Claim 114, wherein the cytotoxic agent is a toxin.

117. The method of Claim 116, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

25 118. The method of Claim 116, wherein the toxin is a maytansinoid.

119. The method of Claim 109, wherein said antibody is produced in bacteria.

120. The method of Claim 109, wherein said antibody is produced in CHO cells.

121. The method of Claim 109, wherein said cell is a cancer cell.

30 122. The method of Claim 121, wherein said cancer cell is further exposed to radiation treatment or a chemotherapeutic agent.

123. The method of Claim 121, wherein said cancer cell is selected from the group consisting of a breast cancer cell, a colorectal cancer cell, a lung cancer cell, an ovarian cancer cell, a central nervous system cancer cell, a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a cervical cancer cell, a melanoma cell and a leukemia cell.

35 124. The method of Claim 123, wherein said protein is more abundantly expressed by said cancer cell as compared to a normal cell of the same tissue origin.

125. The method of Claim 109 which causes the death of said cell.

126. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

127. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for treating a tumor.

5 128. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

129. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

130. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of medicament for treating a tumor.

10 131. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

132. Use of a host cell as claimed in any of Claims 8, 9, 32, or 33 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

15 133. Use of a host cell as claimed in any of Claims 8, 9, 32 or 33 in the preparation of a medicament for treating a tumor.

134. Use of a host cell as claimed in any of Claims 8, 9, 32 or 33 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

135. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

20 136. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for treating a tumor.

137. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

25 138. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

139. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for treating a tumor.

140. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

30 141. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

142. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for treating a tumor.

35 143. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

144. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

145. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for treating a tumor.

146. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

5 147. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

148. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for treating a tumor.

149. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

10 150. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

151. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for treating a tumor.

15 152. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

153. A method for inhibiting the growth of a cell, wherein the growth of said cell is at least in part dependent upon a growth potentiating effect of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

20 (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

25 (e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

30 (f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said protein with an antibody, oligopeptide or organic molecule that binds to said protein, thereby inhibiting the growth of said cell.

154. The method of Claim 153, wherein said cell is a cancer cell.

155. The method of Claim 153, wherein said protein is expressed by said cell.

156. The method of Claim 153, wherein the binding of said antibody, oligopeptide or organic molecule to said protein antagonizes a cell growth-potentiating activity of said protein.

35 157. The method of Claim 153, wherein the binding of said antibody, oligopeptide or organic molecule to said protein induces the death of said cell.

158. The method of Claim 153, wherein said antibody is a monoclonal antibody.

159. The method of Claim 153, wherein said antibody is an antibody fragment.
160. The method of Claim 153, wherein said antibody is a chimeric or a humanized antibody.
161. The method of Claim 153, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.
162. The method of Claim 153, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.
163. The method of Claim 162, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.
164. The method of Claim 162, wherein the cytotoxic agent is a toxin.
165. The method of Claim 164, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.
166. The method of Claim 164, wherein the toxin is a maytansinoid.
167. The method of Claim 153, wherein said antibody is produced in bacteria.
168. The method of Claim 153, wherein said antibody is produced in CHO cells.
169. The method of Claim 153, wherein said protein has:
- (a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);
- (b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;
- (c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;
- (d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;
- (e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or
- (f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).
170. A method of therapeutically treating a tumor in a mammal, wherein the growth of said tumor is at least in part dependent upon a growth potentiating effect of a protein having at least 80% amino acid sequence identity to:
- (a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);
- (b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;
- (c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;
- (d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;
- (e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said protein with an antibody, oligopeptide or organic molecule that binds to said protein, thereby effectively treating said tumor.

171. The method of Claim 170, wherein said protein is expressed by cells of said tumor.

172. The method of Claim 170, wherein the binding of said antibody, oligopeptide or organic molecule to said protein antagonizes a cell growth-potentiating activity of said protein.

173. The method of Claim 170, wherein said antibody is a monoclonal antibody.

174. The method of Claim 170, wherein said antibody is an antibody fragment.

175. The method of Claim 170, wherein said antibody is a chimeric or a humanized antibody.

176. The method of Claim 170, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

177. The method of Claim 170, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

178. The method of Claim 177, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

179. The method of Claim 177, wherein the cytotoxic agent is a toxin.

180. The method of Claim 179, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

181. The method of Claim 179, wherein the toxin is a maytansinoid.

182. The method of Claim 170, wherein said antibody is produced in bacteria.

183. The method of Claim 170, wherein said antibody is produced in CHO cells.

184. The method of Claim 170, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

Yet further embodiments of the present invention will be evident to the skilled artisan upon a reading of the present specification.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a TAT161 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA77507".

Figures 2A-B show a nucleotide sequence (SEQ ID NO:2) of a TAT101 cDNA, wherein SEQ ID NO:2 is a clone designated herein as "DNA80894".

5 Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a TAT157 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA82343".

Figure 4 shows a nucleotide sequence (SEQ ID NO:4) of a TAT160 cDNA, wherein SEQ ID NO:4 is a clone designated herein as "DNA87994".

10 Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a TAT158 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA88131".

Figure 6 shows a nucleotide sequence (SEQ ID NO:6) of a TAT110 cDNA, wherein SEQ ID NO:6 is a clone designated herein as "DNA95930".

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a TAT210 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA95930-1".

15 Figure 8 shows a nucleotide sequence (SEQ ID NO:8) of a TAT159 cDNA, wherein SEQ ID NO:8 is a clone designated herein as "DNA96917".

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a TAT112 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA96930".

20 Figure 10 shows a nucleotide sequence (SEQ ID NO:10) of a TAT147 cDNA, wherein SEQ ID NO:10 is a clone designated herein as "DNA96936".

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a TAT145 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA98565".

Figure 12 shows a nucleotide sequence (SEQ ID NO:12) of a TAT152 cDNA, wherein SEQ ID NO:12 is a clone designated herein as "DNA246435".

25 Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a TAT162 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA98591".

Figure 14 shows a nucleotide sequence (SEQ ID NO:14) of a TAT114 cDNA, wherein SEQ ID NO:14 is a clone designated herein as "DNA108809".

30 Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a TAT119 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA119488".

Figure 16 shows a nucleotide sequence (SEQ ID NO:16) of a TAT103 cDNA, wherein SEQ ID NO:16 is a clone designated herein as "DNA143493".

Figures 17A-B show a nucleotide sequence (SEQ ID NO:17) of a TAT130 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA167234".

35 Figure 18 shows a nucleotide sequence (SEQ ID NO:18) of a TAT166 cDNA, wherein SEQ ID NO:18 is a clone designated herein as "DNA235621".

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a TAT132 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA176766".

Figure 20 shows a nucleotide sequence (SEQ ID NO:20) of a TAT150 cDNA, wherein SEQ ID NO:20 is a clone designated herein as "DNA236463".

5 Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a TAT129 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA181162".

Figure 22 shows a nucleotide sequence (SEQ ID NO:22) of a TAT111 cDNA, wherein SEQ ID NO:22 is a clone designated herein as "DNA188221".

Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a TAT146 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA233876".

10 Figure 24 shows a nucleotide sequence (SEQ ID NO:24) of a TAT148 cDNA, wherein SEQ ID NO:24 is a clone designated herein as "DNA193891".

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a TAT187 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA248170".

15 Figure 26 shows a nucleotide sequence (SEQ ID NO:26) of a TAT118 cDNA, wherein SEQ ID NO:26 is a clone designated herein as "DNA194628".

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a TAT167 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA246415".

Figure 28 shows a nucleotide sequence (SEQ ID NO:28) of a TAT123 cDNA, wherein SEQ ID NO:28 is a clone designated herein as "DNA210499".

20 Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a TAT211 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA219894".

Figure 30 shows a nucleotide sequence (SEQ ID NO:30) of a TAT113 cDNA, wherein SEQ ID NO:30 is a clone designated herein as "DNA215609".

25 Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a TAT128 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA220432".

Figures 32A-B show a nucleotide sequence (SEQ ID NO:32) of a TAT164 cDNA, wherein SEQ ID NO:32 is a clone designated herein as "DNA226094".

Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a TAT122 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA226165".

30 Figure 34 shows a nucleotide sequence (SEQ ID NO:34) of a TAT117 cDNA, wherein SEQ ID NO:34 is a clone designated herein as "DNA226237".

Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a TAT168 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA246450".

35 Figure 36 shows a nucleotide sequence (SEQ ID NO:36) of a TAT144 cDNA, wherein SEQ ID NO:36 is a clone designated herein as "DNA226456".

Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a TAT188 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA237637".



Figure 38 shows a nucleotide sequence (SEQ ID NO:38) of a TAT126 cDNA, wherein SEQ ID NO:38 is a clone designated herein as "DNA226539".

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a TAT151 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA236511".

5 Figure 40 shows a nucleotide sequence (SEQ ID NO:40) of a TAT115 cDNA, wherein SEQ ID NO:40 is a clone designated herein as "DNA226771".

Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a TAT163 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA227087".

Figure 42 shows a nucleotide sequence (SEQ ID NO:42) of a TAT227 cDNA, wherein SEQ ID NO:42 is a clone designated herein as "DNA266307".

10 Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a TAT228 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA266311".

Figure 44 shows a nucleotide sequence (SEQ ID NO:44) of a TAT229 cDNA, wherein SEQ ID NO:44 is a clone designated herein as "DNA266312".

15 Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a TAT230 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA266313".

Figure 46 shows a nucleotide sequence (SEQ ID NO:46) of a TAT121 cDNA, wherein SEQ ID NO:46 is a clone designated herein as "DNA227224".

Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a TAT183 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA247486".

20 Figure 48 shows a nucleotide sequence (SEQ ID NO:48) of a TAT165 cDNA, wherein SEQ ID NO:48 is a clone designated herein as "DNA227578".

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a TAT131 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA227806".

25 Figure 50 shows a nucleotide sequence (SEQ ID NO:50) of a TAT140 cDNA, wherein SEQ ID NO:50 is a clone designated herein as "DNA227904".

Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a TAT127 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA228199".

Figure 52 shows a nucleotide sequence (SEQ ID NO:52) of a TAT116 cDNA, wherein SEQ ID NO:52 is a clone designated herein as "DNA228201".

30 Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a TAT189 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA247488".

Figure 54 shows a nucleotide sequence (SEQ ID NO:54) of a TAT190 cDNA, wherein SEQ ID NO:54 is a clone designated herein as "DNA236538".

35 Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a TAT191 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA247489".

Figure 56 shows a nucleotide sequence (SEQ ID NO:56) of a TAT133 cDNA, wherein SEQ ID NO:56 is a clone designated herein as "DNA228211".

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a TAT186 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA233937".

Figure 58 shows a nucleotide sequence (SEQ ID NO:58) of a TAT120 cDNA, wherein SEQ ID NO:58 is a clone designated herein as "DNA228993".

5 Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a TAT124 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA228994".

Figure 60 shows a nucleotide sequence (SEQ ID NO:60) of a TAT105 cDNA, wherein SEQ ID NO:60 is a clone designated herein as "DNA229410".

Figures 61A-B show a nucleotide sequence (SEQ ID NO:61) of a TAT107 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA229411".

10 Figures 62A-B show a nucleotide sequence (SEQ ID NO:62) of a TAT108 cDNA, wherein SEQ ID NO:62 is a clone designated herein as "DNA229413".

Figures 63A-B show a nucleotide sequence (SEQ ID NO:63) of a TAT139 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA229700".

15 Figure 64 shows a nucleotide sequence (SEQ ID NO:64) of a TAT143 cDNA, wherein SEQ ID NO:64 is a clone designated herein as "DNA231312".

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a TAT100 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA231542".

Figure 66 shows a nucleotide sequence (SEQ ID NO:66) of a TAT284 cDNA, wherein SEQ ID NO:66 is a clone designated herein as "DNA231542-1".

20 Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a TAT285 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA231542-2".

Figure 68 shows a nucleotide sequence (SEQ ID NO:68) of a TAT285-1 cDNA, wherein SEQ ID NO:68 is a clone designated herein as "DNA297393".

25 Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a TAT125 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA232754".

Figure 70 shows a nucleotide sequence (SEQ ID NO:70) of a TAT149 cDNA, wherein SEQ ID NO:70 is a clone designated herein as "DNA234833".

Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a TAT231 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA268022".

30 Figure 72 shows a nucleotide sequence (SEQ ID NO:72) of a TAT153 cDNA, wherein SEQ ID NO:72 is a clone designated herein as "DNA236246".

Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a TAT104 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA236343".

35 Figure 74 shows a nucleotide sequence (SEQ ID NO:74) of a TAT141 cDNA, wherein SEQ ID NO:74 is a clone designated herein as "DNA236493".

Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a TAT102 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA236534".

Figure 76 shows a nucleotide sequence (SEQ ID NO:76) of a TAT109 cDNA, wherein SEQ ID NO:76 is a clone designated herein as "DNA246430".

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a TAT142 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA247480".

5      Figures 78A-B show a nucleotide sequence (SEQ ID NO:78) of a TAT106 cDNA, wherein SEQ ID NO:78 is a clone designated herein as "DNA264454".

Figure 79 shows the amino acid sequence (SEQ ID NO:79) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:2 shown in Figure 2.

10      Figure 81 shows the amino acid sequence (SEQ ID NO:81) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:4 shown in Figure 4.

15      Figure 83 shows the amino acid sequence (SEQ ID NO:83) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:6 shown in Figure 6.

Figure 85 shows the amino acid sequence (SEQ ID NO:85) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

20      Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:8 shown in Figure 8.

Figure 87 shows the amino acid sequence (SEQ ID NO:87) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

25      Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:10 shown in Figure 10.

Figure 89 shows the amino acid sequence (SEQ ID NO:89) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:12 shown in Figure 12.

30      Figure 91 shows the amino acid sequence (SEQ ID NO:91) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:14 shown in Figure 14.

35      Figure 93 shows the amino acid sequence (SEQ ID NO:93) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:16 shown in Figure 16.

Figure 95 shows the amino acid sequence (SEQ ID NO:95) derived from the coding sequence of SEQ ID NO:17 shown in Figures 17A-B.

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:18 shown in Figure 18.

5 Figure 97 shows the amino acid sequence (SEQ ID NO:97) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:20 shown in Figure 20.

Figure 99 shows the amino acid sequence (SEQ ID NO:99) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

10 Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:22 shown in Figure 22.

Figure 101 shows the amino acid sequence (SEQ ID NO:101) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

15 Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:24 shown in Figure 24.

Figure 103 shows the amino acid sequence (SEQ ID NO:103) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:26 shown in Figure 26.

20 Figure 105 shows the amino acid sequence (SEQ ID NO:105) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:28 shown in Figure 28.

25 Figure 107 shows the amino acid sequence (SEQ ID NO:107) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:30 shown in Figure 30.

Figure 109 shows the amino acid sequence (SEQ ID NO:109) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

30 Figures 110A-B shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:32 shown in Figures 32A-B.

Figure 111 shows the amino acid sequence (SEQ ID NO:111) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

35 Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:34 shown in Figure 34.

Figure 113 shows the amino acid sequence (SEQ ID NO:113) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:36 shown in Figure 36.

Figure 115 shows the amino acid sequence (SEQ ID NO:115) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

5 Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:38 shown in Figure 38.

Figure 117 shows the amino acid sequence (SEQ ID NO:117) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:40 shown in Figure 40.

10 Figure 119 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:42 shown in Figure 42.

15 Figure 121 shows the amino acid sequence (SEQ ID NO:121) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:44 shown in Figure 44.

Figure 123 shows the amino acid sequence (SEQ ID NO:123) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

20 Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:46 shown in Figure 46.

Figure 125 shows the amino acid sequence (SEQ ID NO:125) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

25 Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:48 shown in Figure 48.

Figure 127 shows the amino acid sequence (SEQ ID NO:127) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:50 shown in Figure 50.

30 Figure 129 shows the amino acid sequence (SEQ ID NO:129) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:52 shown in Figure 52.

35 Figure 131 shows the amino acid sequence (SEQ ID NO:131) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:54 shown in Figure 54.

Figure 133 shows the amino acid sequence (SEQ ID NO:133) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:56 shown in Figure 56.

5 Figure 135 shows the amino acid sequence (SEQ ID NO:135) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:58 shown in Figure 58.

Figure 137 shows the amino acid sequence (SEQ ID NO:137) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

10 Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:60 shown in Figure 60.

Figure 139 shows the amino acid sequence (SEQ ID NO:139) derived from the coding sequence of SEQ ID NO:61 shown in Figures 61A-B.

15 Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:62 shown in Figures 62A-B.

Figure 141 shows the amino acid sequence (SEQ ID NO:141) derived from the coding sequence of SEQ ID NO:63 shown in Figures 63A-B.

Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:64 shown in Figure 64.

20 Figure 143 shows the amino acid sequence (SEQ ID NO:143) derived from the coding sequence of SEQ ID NO:66 shown in Figure 66.

Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

25 Figure 145 shows the amino acid sequence (SEQ ID NO:145) derived from the coding sequence of SEQ ID NO:68 shown in Figure 68.

Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

Figure 147 shows the amino acid sequence (SEQ ID NO:147) derived from the coding sequence of SEQ ID NO:70 shown in Figure 70.

30 Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

Figure 149 shows the amino acid sequence (SEQ ID NO:149) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

35 Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:74 shown in Figure 74.

Figure 151 shows the amino acid sequence (SEQ ID NO:151) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:76 shown in Figure 76.

Figure 153 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

5 Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:78 shown in Figures 78A-B.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

### I. Definitions

10 The terms "TAT polypeptide" and "TAT" as used herein and when immediately followed by a numerical designation, refer to various polypeptides, wherein the complete designation (i.e., TAT/number) refers to specific polypeptide sequences as described herein. The terms "TAT/number polypeptide" and "TAT/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides, polypeptide variants and fragments of native sequence polypeptides and polypeptide variants (which are further defined herein). The TAT polypeptides described herein may be  
15 isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "TAT polypeptide" refers to each individual TAT/number polypeptide disclosed herein. All disclosures in this specification which refer to the "TAT polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, formation of TAT binding oligopeptides to or against, formation of TAT binding organic molecules to or against, administration of, compositions  
20 containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "TAT polypeptide" also includes variants of the TAT/number polypeptides disclosed herein.

A "native sequence TAT polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding TAT polypeptide derived from nature. Such native sequence TAT polypeptides can be  
25 isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence TAT polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific TAT polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In certain embodiments of the invention, the native sequence TAT polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and  
30 stop codons (if indicated) are shown in bold font and underlined in the figures. Nucleic acid residues indicated as "N" in the accompanying figures are any nucleic acid residue. However, while the TAT polypeptides disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either  
35 upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the TAT polypeptides.

The TAT polypeptide "extracellular domain" or "ECD" refers to a form of the TAT polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a TAT polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the TAT polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a TAT polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various TAT polypeptides disclosed herein may be shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heijne et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"TAT polypeptide variant" means a TAT polypeptide, preferably an active TAT polypeptide, as defined herein having at least about 80% amino acid sequence identity with a full-length native sequence TAT polypeptide sequence as disclosed herein, a TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length TAT polypeptide). Such TAT polypeptide variants include, for instance, TAT polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a TAT polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to a full-length native sequence TAT polypeptide sequence as disclosed herein, a TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide sequence as disclosed herein. Ordinarily, TAT variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240,



250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600 amino acids in length, or more. Optionally, TAT variant polypeptides will have no more than one conservative amino acid substitution as compared to the native TAT polypeptide sequence, alternatively no more than 2, 3, 4, 5, 6, 7, 8, 9, or 10 conservative amino acid substitution as compared to the native TAT polypeptide sequence.

5       "Percent (%) amino acid sequence identity" with respect to the TAT polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific TAT polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid  
10       sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison  
15       computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be  
20       compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

25       In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

30       where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2  
35       and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "TAT", wherein "TAT" represents the amino acid sequence of a hypothetical TAT polypeptide of interest, "Comparison Protein" represents the amino acid

sequence of a polypeptide against which the "TAT" polypeptide of interest is being compared, and "X," "Y" and "Z" each represent different hypothetical amino acid residues. Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

"TAT variant polynucleotide" or "TAT variant nucleic acid sequence" means a nucleic acid molecule which encodes a TAT polypeptide, preferably an active TAT polypeptide, as defined herein and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence TAT polypeptide sequence as disclosed herein, a full-length native sequence TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length TAT polypeptide). Ordinarily, a TAT variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence TAT polypeptide sequence as disclosed herein, a full-length native sequence TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, TAT variant polynucleotides are at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

"Percent (%) nucleic acid sequence identity" with respect to TAT-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the TAT nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user

documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "TAT-DNA", wherein "TAT-DNA" represents a hypothetical TAT-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "TAT-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides. Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

In other embodiments, TAT variant polynucleotides are nucleic acid molecules that encode a TAT polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length TAT polypeptide as disclosed herein. TAT variant polypeptides may be those that are encoded by a TAT variant polynucleotide.

The term "full-length coding region" when used in reference to a nucleic acid encoding a TAT polypeptide refers to the sequence of nucleotides which encode the full-length TAT polypeptide of the invention (which is often shown between start and stop codons, inclusive thereof, in the accompanying figures). The term "full-length coding region" when used in reference to an ATCC deposited nucleic acid refers to the TAT polypeptide-encoding portion of the cDNA that is inserted into the vector deposited with the ATCC (which is often shown between start and stop codons, inclusive thereof, in the accompanying figures).

"Isolated," when used to describe the various TAT polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-

proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the TAT polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" TAT polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Picoil/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) overnight hybridization in a solution that employs 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with a 10 minute wash at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) followed by a 10 minute high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a TAT polypeptide or anti-TAT antibody fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

"Active" or "activity" for the purposes herein refers to form(s) of a TAT polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring TAT, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring TAT other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring TAT and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring TAT.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native TAT polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native TAT polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native TAT polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for

identifying agonists or antagonists of a TAT polypeptide may comprise contacting a TAT polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the TAT polypeptide.

5 "Treating" or "treatment" or "alleviation" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. A subject or mammal is successfully "treated" for a TAT polypeptide-expressing cancer if, after receiving a therapeutic amount of an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule according to the methods of the present invention, the patient shows observable and/or measurable reduction in or absence of one or more of the following: 10 reduction in the number of cancer cells or absence of the cancer cells; reduction in the tumor size; inhibition (i.e., slow to some extent and preferably stop) of cancer cell infiltration into peripheral organs including the spread of cancer into soft tissue and bone; inhibition (i.e., slow to some extent and preferably stop) of tumor metastasis; inhibition, to some extent, of tumor growth; and/or relief to some extent, one or more of the symptoms associated with the specific cancer; reduced morbidity and mortality, and improvement in quality of life issues. To the extent the anti-TAT antibody or TAT binding oligopeptide may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic. Reduction of these signs or symptoms may also be felt by the patient.

The above parameters for assessing successful treatment and improvement in the disease are readily measurable by routine procedures familiar to a physician. For cancer therapy, efficacy can be measured, for 20 example, by assessing the time to disease progression (TTP) and/or determining the response rate (RR). Metastasis can be determined by staging tests and by bone scan and tests for calcium level and other enzymes to determine spread to the bone. CT scans can also be done to look for spread to the pelvis and lymph nodes in the area. Chest X-rays and measurement of liver enzyme levels by known methods are used to look for metastasis to the lungs and liver, respectively. Other routine methods for monitoring the disease include transrectal ultrasonography (TRUS) and transrectal needle biopsy (TRNB).

25 For bladder cancer, which is a more localized cancer, methods to determine progress of disease include urinary cytologic evaluation by cystoscopy, monitoring for presence of blood in the urine, visualization of the urothelial tract by sonography or an intravenous pyelogram, computed tomography (CT) and magnetic resonance imaging (MRI). The presence of distant metastases can be assessed by CT of the abdomen, chest x-rays, or radionuclide imaging of the skeleton.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

35 "Mammal" for purposes of the treatment of, alleviating the symptoms of or diagnosis of a cancer refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is

human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN®, polyethylene glycol (PEG), and PLURONICS®.

By "solid phase" or "solid support" is meant a non-aqueous matrix to which an antibody, TAT binding oligopeptide or TAT binding organic molecule of the present invention can adhere or attach. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a TAT polypeptide, an antibody thereto or a TAT binding oligopeptide) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small" molecule or "small" organic molecule is defined herein to have a molecular weight below about 500 Daltons.

An "effective amount" of a polypeptide, antibody, TAT binding oligopeptide, TAT binding organic molecule or an agonist or antagonist thereof as disclosed herein is an amount sufficient to carry out a specifically stated purpose. An "effective amount" may be determined empirically and in a routine manner, in relation to the stated purpose.

The term "therapeutically effective amount" refers to an amount of an antibody, polypeptide, TAT binding oligopeptide, TAT binding organic molecule or other drug effective to "treat" a disease or disorder in a subject or mammal. In the case of cancer, the therapeutically effective amount of the drug may reduce the number of cancer cells; reduce the tumor size; inhibit (i.e., slow to some extent and preferably stop) cancer cell infiltration into peripheral organs; inhibit (i.e., slow to some extent and preferably stop) tumor metastasis; inhibit, to some extent, tumor growth; and/or relieve to some extent one or more of the symptoms associated with the cancer. See the definition herein of "treating". To the extent the drug may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic.

A "growth inhibitory amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide or TAT binding organic molecule is an amount capable of inhibiting the growth of a cell, especially tumor, e.g., cancer cell, either *in vitro* or *in vivo*. A "growth inhibitory amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide or TAT binding organic molecule for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

5 A "cytotoxic amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide or TAT binding organic molecule is an amount capable of causing the destruction of a cell, especially tumor, e.g., cancer cell, either *in vitro* or *in vivo*. A "cytotoxic amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide or TAT binding organic molecule for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

10 The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-TAT monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-TAT antibody compositions with polypeptidic specificity, polyclonal antibodies, single chain anti-TAT antibodies, and fragments of anti-TAT antibodies (see below) as long as they exhibit the desired biological or immunological activity. The term "immunoglobulin" (Ig) is used interchangeable with antibody herein.

15 An "isolated antibody" is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequencer, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

25 The basic 4-chain antibody unit is a heterotetrameric glycoprotein composed of two identical light (L) chains and two identical heavy (H) chains (an IgM antibody consists of 5 of the basic heterotetramer unit along with an additional polypeptide called J chain, and therefore contain 10 antigen binding sites, while secreted IgA antibodies can polymerize to form polyvalent assemblages comprising 2-5 of the basic 4-chain units along with J chain). In the case of IgG, the 4-chain unit is generally about 150,000 daltons. Each L chain is linked to a H chain by one covalent disulfide bond, while the two H chains are linked to each other by one or more disulfide bonds depending on the H chain isotype. Each H and L chain also has regularly spaced intrachain disulfide bridges. Each H chain has at the N-terminus, a variable domain ( $V_H$ ) followed by three constant domains ( $C_H$ ) for each of the  $\alpha$  and  $\gamma$  chains and four  $C_H$  domains for  $\mu$  and  $\epsilon$  isotypes. Each L chain has at the N-terminus, a variable domain ( $V_L$ ) followed by a constant domain ( $C_L$ ) at its other end. The  $V_L$  is aligned with the  $V_H$  and the  $C_L$  is aligned with the first constant domain of the heavy chain ( $C_H1$ ). Particular amino acid residues are believed to form an interface between the light chain and heavy chain variable domains. The pairing of a  $V_H$  and  $V_L$  together forms a single antigen-binding site. For the structure and properties of the



different classes of antibodies, see, e.g., Basic and Clinical Immunology, 8th edition, Daniel P. Stites, Abba I. Terr and Tristram G. Parslow (eds.), Appleton & Lange, Norwalk, CT, 1994, page 71 and Chapter 6.

The L chain from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains. Depending on the amino acid sequence of the constant domain of their heavy chains ( $C_H$ ), immunoglobulins can be assigned to different classes or isotypes. There are five classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, having heavy chains designated  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ , and  $\mu$ , respectively. The  $\gamma$  and  $\alpha$  classes are further divided into subclasses on the basis of relatively minor differences in  $C_H$  sequence and function, e.g., humans express the following subclasses: IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2.

The term "variable" refers to the fact that certain segments of the variable domains differ extensively in sequence among antibodies. The V domain mediates antigen binding and define specificity of a particular antibody for its particular antigen. However, the variability is not evenly distributed across the 110-amino acid span of the variable domains. Instead, the V regions consist of relatively invariant stretches called framework regions (FRs) of 15-30 amino acids separated by shorter regions of extreme variability called "hypervariable regions" that are each 9-12 amino acids long. The variable domains of native heavy and light chains each comprise four FRs, largely adopting a  $\beta$ -sheet configuration, connected by three hypervariable regions, which form loops connecting, and in some cases forming part of, the  $\beta$ -sheet structure. The hypervariable regions in each chain are held together in close proximity by the FRs and, with the hypervariable regions from the other chain, contribute to the formation of the antigen-binding site of antibodies (see Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD, (1991)). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody dependent cellular cytotoxicity (ADCC).

The term "hypervariable region" when used herein refers to the amino acid residues of an antibody which are responsible for antigen-binding. The hypervariable region generally comprises amino acid residues from a "complementarity determining region" or "CDR" (e.g. around about residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the  $V_L$ , and around about 1-35 (H1), 50-65 (H2) and 95-102 (H3) in the  $V_H$ ; Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD, (1991)) and/or those residues from a "hypervariable loop" (e.g. residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the  $V_L$ , and 26-32 (H1), 53-55 (H2) and 96-101 (H3) in the  $V_H$ ; Chothia and Lesk J. Mol. Biol., 196:901-917 (1987)).

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations which include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they may be synthesized uncontaminated

by other antibodies. The modifier "monoclonal" is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies useful in the present invention may be prepared by the hybridoma methodology first described by Kohler et al., *Nature*, 256:495 (1975), or may be made using recombinant DNA methods in bacterial, eukaryotic animal or plant cells (see, e.g., U.S. Patent No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using the techniques described in Clackson et al., *Nature*, 352:624-628 (1991) and Marks et al., *J. Mol. Biol.*, 222:581-597 (1991), for example.

The monoclonal antibodies herein include "chimeric" antibodies in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (see U.S. Patent No. 4,816,567; and Morrison et al., *Proc. Natl. Acad. Sci. USA*, 81:6851-6855 (1984)). Chimeric antibodies of interest herein include "primatized" antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (e.g. Old World Monkey, Ape etc), and human constant region sequences.

An "intact" antibody is one which comprises an antigen-binding site as well as a C<sub>L</sub> and at least heavy chain constant domains, C<sub>H</sub>1, C<sub>H</sub>2 and C<sub>H</sub>3. The constant domains may be native sequence constant domains (e.g. human native sequence constant domains) or amino acid sequence variant thereof. Preferably, the intact antibody has one or more effector functions.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')<sub>2</sub>, and Fv fragments; diabodies; linear antibodies (see U.S. Patent No. 5,641,870, Example 2; Zapata et al., *Protein Eng.*, 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. The Fab fragment consists of an entire L chain along with the variable region domain of the H chain (V<sub>H</sub>), and the first constant domain of one heavy chain (C<sub>H</sub>1). Each Fab fragment is monovalent with respect to antigen binding, i.e., it has a single antigen-binding site. Pepsin treatment of an antibody yields a single large F(ab')<sub>2</sub> fragment which roughly corresponds to two disulfide linked Fab fragments having divalent antigen-binding activity and is still capable of cross-linking antigen. Fab' fragments differ from Fab fragments by having additional few residues at the carboxy terminus of the C<sub>H</sub>1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')<sub>2</sub> antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The Fc fragment comprises the carboxy-terminal portions of both H chains held together by disulfides. The effector functions of antibodies are determined by sequences in the Fc region, which region is also the part

recognized by Fc receptors (FcR) found on certain types of cells.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This fragment consists of a dimer of one heavy- and one light-chain variable region domain in tight, non-covalent association. From the folding of these two domains emanate six hypervariable loops (3 loops each from the H and L chain) that contribute the amino acid residues for antigen binding and confer antigen binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

"Single-chain Fv" also abbreviated as "sFv" or "scFv" are antibody fragments that comprise the V<sub>H</sub> and V<sub>L</sub> antibody domains connected into a single polypeptide chain. Preferably, the sFv polypeptide further comprises a polypeptide linker between the V<sub>H</sub> and V<sub>L</sub> domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Packthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994); Borrebueck 1995, *infra*.

The term "diabodies" refers to small antibody fragments prepared by constructing sFv fragments (see preceding paragraph) with short linkers (about 5-10 residues) between the V<sub>H</sub> and V<sub>L</sub> domains such that inter-chain but not intra-chain pairing of the V domains is achieved, resulting in a bivalent fragment, i.e., fragment having two antigen-binding sites. Bispecific diabodies are heterodimers of two "crossover" sFv fragments in which the V<sub>H</sub> and V<sub>L</sub> domains of the two antibodies are present on different polypeptide chains. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., *Proc. Natl. Acad. Sci. USA*, 90:6444-6448 (1993).

"Humanized" forms of non-human (e.g., rodent) antibodies are chimeric antibodies that contain minimal sequence derived from the non-human antibody. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hypervariable region of the recipient are replaced by residues from a hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit or non-human primate having the desired antibody specificity, affinity, and capability. In some instances, framework region (FR) residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops correspond to those of a non-human immunoglobulin and all or substantially all of the FRs are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones et al., *Nature* 321:522-525 (1986); Riechmann et al., *Nature* 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992).

A "species-dependent antibody," e.g., a mammalian anti-human IgE antibody, is an antibody which has a stronger binding affinity for an antigen from a first mammalian species than it has for a homologue of that antigen from a second mammalian species. Normally, the species-dependent antibody "bind specifically"

to a human antigen (i.e., has a binding affinity (Kd) value of no more than about  $1 \times 10^{-7}$  M, preferably no more than about  $1 \times 10^{-8}$  and most preferably no more than about  $1 \times 10^{-9}$  M) but has a binding affinity for a homologue of the antigen from a second non-human mammalian species which is at least about 50 fold, or at least about 500 fold, or at least about 1000 fold, weaker than its binding affinity for the human antigen. The species-dependent antibody can be of any of the various types of antibodies as defined above, but preferably is a humanized or human antibody.

A "TAT binding oligopeptide" is an oligopeptide that binds, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. TAT binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, e.g., U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen et al., in Synthetic Peptides as Antigens, 130-149 (1986); Geysen et al., J. Immunol. Meth., 102:259-274 (1987); Schoofs et al., J. Immunol., 140:611-616 (1988), Cwiria, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clarkson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol., 222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363, and Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

A "TAT binding organic molecule" is an organic molecule other than an oligopeptide or antibody as defined herein that binds, preferably specifically, to a TAT polypeptide as described herein. TAT binding organic molecules may be identified and chemically synthesized using known methodology (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules are usually less than about 2000 daltons in size, alternatively less than about 1500, 750, 500, 250 or 200 daltons in size, wherein such organic molecules that are capable of binding, preferably specifically, to a TAT polypeptide as described herein may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening organic molecule libraries for molecules that are capable of binding to a polypeptide target are well known in the art (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585).

An antibody, oligopeptide or other organic molecule "which binds" an antigen of interest, e.g. a tumor-associated polypeptide antigen target, is one that binds the antigen with sufficient affinity such that the antibody, oligopeptide or other organic molecule is useful as a diagnostic and/or therapeutic agent in targeting

a cell or tissue expressing the antigen, and does not significantly cross-react with other proteins. In such embodiments, the extent of binding of the antibody, oligopeptide or other organic molecule to a "non-target" protein will be less than about 10% of the binding of the antibody, oligopeptide or other organic molecule to its particular target protein as determined by fluorescence activated cell sorting (FACS) analysis or radioimmuno-precipitation (RIA). With regard to the binding of an antibody, oligopeptide or other organic molecule to a target molecule, the term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target means binding that is measurably different from a non-specific interaction. Specific binding can be measured, for example, by determining binding of a molecule compared to binding of a control molecule, which generally is a molecule of similar structure that does not have binding activity. For example, specific binding can be determined by competition with a control molecule that is similar to the target, for example, an excess of non-labeled target. In this case, specific binding is indicated if the binding of the labeled target to a probe is competitively inhibited by excess unlabeled target. The term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target as used herein can be exhibited, for example, by a molecule having a Kd for the target of at least about  $10^{-6}$  M, alternatively at least about  $10^{-5}$  M, alternatively at least about  $10^{-4}$  M, alternatively at least about  $10^{-3}$  M, alternatively at least about  $10^{-2}$  M, alternatively at least about  $10^{-10}$  M, alternatively at least about  $10^{-11}$  M, alternatively at least about  $10^{-12}$  M, or greater. In one embodiment, the term "specific binding" refers to binding where a molecule binds to a particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

An antibody, oligopeptide or other organic molecule that "inhibits the growth of tumor cells expressing a TAT polypeptide" or a "growth inhibitory" antibody, oligopeptide or other organic molecule is one which results in measurable growth inhibition of cancer cells expressing or overexpressing the appropriate TAT polypeptide. The TAT polypeptide may be a transmembrane polypeptide expressed on the surface of a cancer cell or may be a polypeptide that is produced and secreted by a cancer cell. Preferred growth inhibitory anti-TAT antibodies, oligopeptides or organic molecules inhibit growth of TAT-expressing tumor cells by greater than 20%, preferably from about 20% to about 50%, and even more preferably, by greater than 50% (e.g., from about 50% to about 100%) as compared to the appropriate control, the control typically being tumor cells not treated with the antibody, oligopeptide or other organic molecule being tested. In one embodiment, growth inhibition can be measured at an antibody concentration of about 0.1 to 30  $\mu$ g/ml or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. Growth inhibition of tumor cells *in vivo* can be determined in various ways such as is described in the Experimental Examples section below. The antibody is growth inhibitory *in vivo* if administration of the anti-TAT antibody at about 1  $\mu$ g/kg to about 100 mg/kg body weight results in reduction in tumor size or tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

An antibody, oligopeptide or other organic molecule which "induces apoptosis" is one which induces programmed cell death as determined by binding of annexin V, fragmentation of DNA, cell shrinkage, dilation

of endoplasmic reticulum, cell fragmentation, and/or formation of membrane vesicles (called apoptotic bodies). The cell is usually one which overexpresses a TAT polypeptide. Preferably the cell is a tumor cell, e.g., a prostate, breast, ovarian, stomach, endometrial, lung, kidney, colon, bladder cell. Various methods are available for evaluating the cellular events associated with apoptosis. For example, phosphatidyl serine (PS) translocation can be measured by annexin binding; DNA fragmentation can be evaluated through DNA laddering; and nuclear/chromatin condensation along with DNA fragmentation can be evaluated by any increase in hypodiploid cells. Preferably, the antibody, oligopeptide or other organic molecule which induces apoptosis is one which results in about 2 to 50 fold, preferably about 5 to 50 fold, and most preferably about 10 to 50 fold, induction of annexin binding relative to untreated cell in an annexin binding assay.

Antibody "effector functions" refer to those biological activities attributable to the Fc region (a native sequence Fc region or amino acid sequence variant Fc region) of an antibody, and vary with the antibody isotype. Examples of antibody effector functions include: C1q binding and complement dependent cytotoxicity; Fc receptor binding; antibody-dependent cell-mediated cytotoxicity (ADCC); phagocytosis; down regulation of cell surface receptors (e.g., B cell receptor); and B cell activation.

"Antibody-dependent cell-mediated cytotoxicity" or "ADCC" refers to a form of cytotoxicity in which secreted Ig bound onto Fc receptors (FcRs) present on certain cytotoxic cells (e.g., Natural Killer (NK) cells, neutrophils, and macrophages) enable these cytotoxic effector cells to bind specifically to an antigen-bearing target cell and subsequently kill the target cell with cytotoxins. The antibodies "arm" the cytotoxic cells and are absolutely required for such killing. The primary cells for mediating ADCC, NK cells, express FcγRIII only, whereas monocytes express FcγRI, FcγRII and FcγRIII. FcR expression on hematopoietic cells is summarized in Table 3 on page 464 of Ravetch and Kinet, Annu. Rev. Immunol. 9:457-92 (1991). To assess ADCC activity of a molecule of interest, an in vitro ADCC assay, such as that described in US Patent No. 5,500,362 or 5,821,337 may be performed. Useful effector cells for such assays include peripheral blood mononuclear cells (PBMC) and Natural Killer (NK) cells. Alternatively, or additionally, ADCC activity of the molecule of interest may be assessed in vivo, e.g., in a animal model such as that disclosed in Clynes et al. (USA) 95:652-656 (1998).

"Fc receptor" or "FcR" describes a receptor that binds to the Fc region of an antibody. The preferred FcR is a native sequence human FcR. Moreover, a preferred FcR is one which binds an IgG antibody (a gamma receptor) and includes receptors of the FcγRI, FcγRII and FcγRIII subclasses, including allelic variants and alternatively spliced forms of these receptors. FcγRII receptors include FcγRIIA (an "activating receptor") and FcγRIIB (an "inhibiting receptor"), which have similar amino acid sequences that differ primarily in the cytoplasmic domains thereof. Activating receptor FcγRIIA contains an immunoreceptor tyrosine-based activation motif (ITAM) in its cytoplasmic domain. Inhibiting receptor FcγRIIB contains an immunoreceptor tyrosine-based inhibition motif (ITIM) in its cytoplasmic domain. (see review M. in Daëron, Annu. Rev. Immunol. 15:203-234 (1997)). FcRs are reviewed in Ravetch and Kinet, Annu. Rev. Immunol. 9:457-492 (1991); Capei et al., Immunomethods 4:25-34 (1994); and de Haas et al., J. Lab. Clin. Med. 126:330-41 (1995). Other FcRs, including those to be identified in the future, are encompassed by the term "FcR" herein. The term also includes the neonatal receptor, FcRn, which is responsible for the transfer of

maternal IgGs to the fetus (Guyer et al., J. Immunol. 117:587 (1976) and Kim et al., J. Immunol. 24:249 (1994)).

"Human effector cells" are leukocytes which express one or more FcRs and perform effector functions. Preferably, the cells express at least FcγRIII and perform ADCC effector function. Examples of human leukocytes which mediate ADCC include peripheral blood mononuclear cells (PBMC), natural killer (NK) cells, monocytes, cytotoxic T cells and neutrophils; with PBMCs and NK cells being preferred. The effector cells may be isolated from a native source, e.g., from blood.

"Complement dependent cytotoxicity" or "CDC" refers to the lysis of a target cell in the presence of complement. Activation of the classical complement pathway is initiated by the binding of the first component of the complement system (C1q) to antibodies (of the appropriate subclass) which are bound to their cognate antigen. To assess complement activation, a CDC assay, e.g., as described in Gazzano-Santoro et al., J. Immunol. Methods 202:163 (1996), may be performed.

The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include, but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia or lymphoid malignancies. More particular examples of such cancers include squamous cell cancer (e.g., epithelial squamous cell cancer), lung cancer including small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung and squamous carcinoma of the lung, cancer of the peritoneum, hepatocellular cancer, gastric or stomach cancer including gastrointestinal cancer, pancreatic cancer, glioblastoma, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, hepatoma, breast cancer, colon cancer, rectal cancer, colorectal cancer, endometrial or uterine carcinoma, salivary gland carcinoma, kidney or renal cancer, prostate cancer, vulval cancer, thyroid cancer, hepatic carcinoma, anal carcinoma, penile carcinoma, melanoma, multiple myeloma and B-cell lymphoma, brain, as well as head and neck cancer, and associated metastases.

The terms "cell proliferative disorder" and "proliferative disorder" refer to disorders that are associated with some degree of abnormal cell proliferation. In one embodiment, the cell proliferative disorder is cancer.

"Tumor", as used herein, refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

An antibody, oligopeptide or other organic molecule which "induces cell death" is one which causes a viable cell to become nonviable. The cell is one which expresses a TAT polypeptide, preferably a cell that overexpresses a TAT polypeptide as compared to a normal cell of the same tissue type. The TAT polypeptide may be a transmembrane polypeptide expressed on the surface of a cancer cell or may be a polypeptide that is produced and secreted by a cancer cell. Preferably, the cell is a cancer cell, e.g., a breast, ovarian, stomach, endometrial, salivary gland, lung, kidney, colon, thyroid, pancreatic or bladder cell. Cell death *in vitro* may be determined in the absence of complement and immune effector cells to distinguish cell death induced by antibody-dependent cell-mediated cytotoxicity (ADCC) or complement dependent cytotoxicity (CDC). Thus, the assay for cell death may be performed using heat inactivated serum (i.e., in the absence of complement) and in the absence of immune effector cells. To determine whether the antibody, oligopeptide

or other organic molecule is able to induce cell death, loss of membrane integrity as evaluated by uptake of propidium iodide (PI), trypan blue (see Moore et al. *Cytotechnology* 17:1-11 (1995)) or 7AAD can be assessed relative to untreated cells. Preferred cell death-inducing antibodies, oligopeptides or other organic molecules are those which induce PI uptake in the PI uptake assay in BT474 cells.

A "TAT-expressing cell" is a cell which expresses an endogenous or transfected TAT polypeptide either on the cell surface or in a secreted form. A "TAT-expressing cancer" is a cancer comprising cells that have a TAT polypeptide present on the cell surface or that produce and secrete a TAT polypeptide. A "TAT-expressing cancer" optionally produces sufficient levels of TAT polypeptide on the surface of cells thereof, such that an anti-TAT antibody, oligopeptide or other organic molecule can bind thereto and have a therapeutic effect with respect to the cancer. In another embodiment, a "TAT-expressing cancer" optionally produces and secretes sufficient levels of TAT polypeptide, such that an anti-TAT antibody, oligopeptide or other organic molecule antagonist can bind thereto and have a therapeutic effect with respect to the cancer. With regard to the latter, the antagonist may be an antisense oligonucleotide which reduces, inhibits or prevents production and secretion of the secreted TAT polypeptide by tumor cells. A cancer which "overexpresses" a TAT polypeptide is one which has significantly higher levels of TAT polypeptide at the cell surface thereof, or produces and secretes, compared to a noncancerous cell of the same tissue type. Such overexpression may be caused by gene amplification or by increased transcription or translation. TAT polypeptide overexpression may be determined in a diagnostic or prognostic assay by evaluating increased levels of the TAT protein present on the surface of a cell, or secreted by the cell (e.g., via an immunohistochemistry assay using anti-TAT antibodies prepared against an isolated TAT polypeptide which may be prepared using recombinant DNA technology from an isolated nucleic acid encoding the TAT polypeptide; FACS analysis, etc.). Alternatively, or additionally, one may measure levels of TAT polypeptide-encoding nucleic acid or mRNA in the cell, e.g., via fluorescent *in situ* hybridization using a nucleic acid based probe corresponding to a TAT-encoding nucleic acid or the complement thereof; (FISH; see WO98/45479 published October, 1998), Southern blotting, Northern blotting, or polymerase chain reaction (PCR) techniques, such as real time quantitative PCR (RT-PCR). One may also study TAT polypeptide overexpression by measuring shed antigen in a biological fluid such as serum, e.g., using antibody-based assays (see also, e.g., U.S. Patent No. 4,933,294 issued June 12, 1990; WO91/05264 published April 18, 1991; U.S. Patent 5,461,638 issued March 28, 1995; and Sias et al., *J. Immunol. Methods* 132:73-83 (1990)). Aside from the above assays, various *in vivo* assays are available to the skilled practitioner. For example, one may expose cells within the body of the patient to an antibody which is optionally labeled with a detectable label, e.g., a radioactive isotope, and binding of the antibody to cells in the patient can be evaluated, e.g., by external scanning for radioactivity or by analyzing a biopsy taken from a patient previously exposed to the antibody.

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin



molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody, oligopeptide or other organic molecule so as to generate a "labeled" antibody, oligopeptide or other organic molecule. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (e.g.,  $^{211}\text{At}$ ,  $^{131}\text{I}$ ,  $^{125}\text{I}$ ,  $^{90}\text{Y}$ ,  $^{186}\text{Re}$ ,  $^{188}\text{Re}$ ,  $^{153}\text{Sm}$ ,  $^{212}\text{Bi}$ ,  $^{32}\text{P}$  and radioactive isotopes of Lu), chemotherapeutic agents e.g. methotrexate, adriamycin, vinca alkaloids (vincristine, vinblastine, etoposide), doxorubicin, melphalan, mitomycin C, chlorambucil, daunorubicin or other intercalating agents, enzymes and fragments thereof such as nucleolytic enzymes, antibiotics, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof, and the various antitumor or anticancer agents disclosed below. Other cytotoxic agents are described below. A tumoricidal agent causes destruction of tumor cells.

A "growth inhibitory agent" when used herein refers to a compound or composition which inhibits growth of a cell, especially a TAT-expressing cancer cell, either *in vitro* or *in vivo*. Thus, the growth inhibitory agent may be one which significantly reduces the percentage of TAT-expressing cells in S phase. Examples of growth inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxanes, and topoisomerase II inhibitors such as doxorubicin, epirubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further information can be found in The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogenes, and antineoplastic drugs" by Murakami et al. (WB Saunders: Philadelphia, 1995), especially p. 13. The taxanes (paclitaxel and docetaxel) are anticancer drugs both derived from the yew tree. Docetaxel (TAXOTERE®, Rhone-Poulenc Rorer), derived from the European yew, is a semisynthetic analogue of paclitaxel (TAXOL®, Bristol-Myers Squibb). Paclitaxel and docetaxel promote the assembly of microtubules from tubulin dimers and stabilize microtubules by preventing depolymerization, which results in the inhibition of mitosis in cells.

"Doxorubicin" is an anthracycline antibiotic. The full chemical name of doxorubicin is (3S-cis)-10-[(3-amino-2,3,6-trideoxy- $\alpha$ -L-xylo-hexapyranosyl)oxy]-7,8,9,10-tetrahydro-6,8,11-trihydroxy-8-(hydroxyacetyl)-1-methoxy-5,12-naphthacenedione.

The term "cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and

traditional polypeptide hormones. Included among the cytokines are growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); hepatic growth factor; fibroblast growth factor; prolactin; placental lactogen; tumor necrosis factor- $\alpha$  and - $\beta$ ; muellerian-inhibiting substance; mouse gonadotropin-associated peptide; inhibin; activin; vascular endothelial growth factor; integrin; thrombopoietin (TPO); nerve growth factors such as NGF- $\beta$ ; platelet-growth factor; transforming growth factors (TGFs) such as TGF- $\alpha$  and TGF- $\beta$ ; insulin-like growth factor-I and -II; erythropoietin (EPO); osteoinductive factors; interferons such as interferon - $\alpha$ , - $\beta$ , and - $\gamma$ ; colony stimulating factors (CSFs) such as macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); interleukins (ILs) such as IL-1, IL-1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; a tumor necrosis factor such as TNF- $\alpha$  or TNF- $\beta$ ; and other polypeptide factors including LIF and kit ligand (KL). As used herein, the term cytokine includes proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence cytokines.

The term "package insert" is used to refer to instructions customarily included in commercial packages of therapeutic products, that contain information about the indications, usage, dosage, administration, contraindications and/or warnings concerning the use of such therapeutic products.

**Table 1**

```

/*
*
* C-C increased from 12 to 15
* Z is average of EQ
5  * B is average of ND
* nmuch with stop is _M; stop-stop = 0; J (joker) match = 0
*/
#define _M -8 /* value of a match with a stop */

10 int _day[26][26] = {
/* A */ {2, 0, -2, 0, 0, -4, 1, -1, 1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ {0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, 4, _M, -3, -5, 4, 0, -2, 0, -2, -8, 0, 0, 5},
15 /* D */ {0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ {0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, 3, 0, -1, 0, 0, 7, -5},
/* G */ {1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, 1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, 1, 0, -2, -3, 0, 0, 2},
20 /* I */ {1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, 1},
25 /* N */ {0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ {_M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ {1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, 1, -6, 0, -5, 0},
/* Q */ {0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
30 /* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ {1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ {1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 5, 0, 0, -5, 0, -3, 0},
/* U */ {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ {0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
35 /* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -3, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ {0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

Table 1 (cont')

```

/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP 16 /* max jumps in a diag */
#define MAXGAP 24 /* don't continue to penalize gaps larger than this */
#define JMPS 1024 /* max jumps in a path */
#define MX 4 /* save if there's at least MX-1 bases since last jump */

#define DMAT 3 /* value of matching bases */
#define DMIS 0 /* penalty for mismatched bases */
#define DINS0 8 /* penalty for a gap */
#define DINS1 1 /* penalty per base */
#define PINS0 8 /* penalty for a gap */
#define PINS1 4 /* penalty per residue */

struct jmp {
    short n[MAXJMP]; /* size of jump (neg for del) */
    unsigned short x[MAXJMP]; /* base no. of jump in seq x */
}; /* limits seq to 2^16 -1 */

struct diag {
    int score; /* score at last jmp */
    long offset; /* offset of prev block */
    short jmp; /* current jmp index */
    struct jmp jp; /* list of jumps */
};

struct path {
    int spc; /* number of leading spaces */
    short n[JMP]; /* size of jmp (gap) */
    int x[JMP]; /* loc of jmp (last elem before gap) */
};

char *ofile; /* output file name */
char *names[2]; /* seq names; getseqs() */
char *prog; /* prog name for err msgs */
char *seqs[2]; /* seqs; getseqs() */
int dmax; /* test diag: nw() */
int dmax0; /* final diag */
int dma; /* set if dma: main() */
int endgaps; /* set if penalizing end gaps */
int gapc, gapy; /* total gaps in seqs */
int len0, len1; /* seq lens */
int ngapc, ngapy; /* total size of gaps */
int snax; /* max score: nw() */
int *xbm; /* bitmap for matching */
long offset; /* current offset in jmp file */
struct diag dx; /* holds diagonals */
struct path pp[2]; /* holds path for seqs */

char *calloc(), *malloc(), *index(), *strcpy();
char *getseq(), *g_calloc();

```

Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: prog file1 file2
* where file1 and file2 are two dna or two protein sequences.
5 * The sequences can be in upper- or lower-case an may contain ambiguity
* Any lines beginning with '!', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
10 * Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nv.h"
15 #include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,13,0,0,0,5,6,8,8,7,9,0,10,0
};

20 static _pbval[26] = {
    1, 2, {1 < < ('D'-'A')}, {1 < < ('N'-'A')}, 4, 8, 16, 32, 64,
    128, 256, 0xFF, 0xFF, 1 < < 10, 1 < < 11, 1 < < 12, 1 < < 13, 1 < < 14,
    1 < < 15, 1 < < 16, 1 < < 17, 1 < < 18, 1 < < 19, 1 < < 20, 1 < < 21, 1 < < 22,
25 1 < < 23, 1 < < 24, 1 < < 25, {1 < < ('E'-'A')}, {1 < < ('Q'-'A')}
};

main(ac, av)
30     int    ac;
     char    *av[];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with '!', '>' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
40     }
    names[0] = av[1];
    names[1] = av[2];
    seqx[0] = getseq(names[0], &len0);
    seqx[1] = getseq(names[1], &len1);
45     xbm = (dms)? _dbval : _pbval;

    endgaps = 0; /* 1 to penalize endgaps */
    ofile = "align.out"; /* output file */

50     nwq; /* fill in the matrix, get the possible jmps */
    readjmps(); /* get the actual jmps */
    print(); /* print stats, alignment */

    cleanup(); /* unlink any tmp files */
55 }

```

**Table 1 (cont')**

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seq y
 * to a gap in seq x.
 */
nw()
{
    char      *px, *py;           /* seqs and ptrs */
    int      *ndely, *dely;       /* keep track of dely */
    int      nidx, delx;          /* keep track of delx */
    int      *trup;               /* for swapping row0, row1 */
    int      mis;                 /* score for each type */
    int      ins0, ins1;          /* insertion penalties */
    register  id;                 /* diagonal index */
    register  lj;                 /* jmp index */
    register  *col0, *col1;       /* score for curr, last row */
    register  xx, yy;             /* index into seqs */

    dx = (struct diag *)g_calloc("to get diag", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0; /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

HW

Table 1 (cont')

...HW

```

for (py = seqx[1], yy = 1; yy <= lcal; py++, yy++) {
    mis = colQ[yy-1];
    if (dna)
        mis += (xbm["px-'A'"]&xbm["py-'A'"])? DMAT : DMIS;
    else
        mis += _day["px-'A'"]["py-'A'"];

    /* update penalty for del in x seq;
     * favor new del over ongoing del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (colQ[yy] - ins0 >= dely[yy]) {
            dely[yy] = colQ[yy] - (ins0 + ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] += ins1;
            ndely[yy]++;
        }
    } else {
        if (colQ[yy] - (ins0 + ins1) >= dely[yy]) {
            dely[yy] = colQ[yy] - (ins0 + ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongoing del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (colI[yy-1] - ins0 >= delx) {
            delx = colI[yy-1] - (ins0 + ins1);
            ndelx = 1;
        } else {
            delx += ins1;
            ndelx++;
        }
    } else {
        if (colI[yy-1] - (ins0 + ins1) >= delx) {
            delx = colI[yy-1] - (ins0 + ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */

```

Table 1 (cont')

...NW

```

id = xx - yy + len1 - 1;
if (mis >= dclx && mis >= dely[yy])
    coll[yy] = mis;
else if (dclx >= dely[yy]) {
    coll[yy] = dclx;
    ij = dx[id].jmp;
    if (dx[id].jp.n[0] && (idna || (ndclx >= MAXIMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINSO)) {
        dx[id].jmp++;
        if (++ij >= MAXIMP) {
            writeimps(id);
            ij = dx[id].jmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = ndclx;
        dx[id].jp.x[ij] = xx;
        dx[id].score = dclx;
    }
} else {
    coll[yy] = dely[yy];
    ij = dx[id].jmp;
    if (dx[id].jp.n[0] && (idna || (ndely[yy] >= MAXIMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINSO)) {
        dx[id].jmp++;
        if (++ij >= MAXIMP) {
            writeimps(id);
            ij = dx[id].jmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = ndely[yy];
        dx[id].jp.x[ij] = xx;
        dx[id].score = dely[yy];
    }
}
if (xx == len0 && yy < len1) {
    /* last col
    */
    if (endgaps)
        coll[yy] := ins0+ins1*(len1-yy);
    if (coll[yy] > smax) {
        smax = coll[yy];
        dmax = id;
    }
}
}
if (endgaps && xx < len0)
    coll[yy-1] := ins0+ins1*(len0-xx);
if (coll[yy-1] > smax) {
    smax = coll[yy-1];
    dmax = id;
}
}
jmp = col0; col0 = coll; coll = jmp;
}
(void) free((char *)ndclx);
(void) free((char *)ndely);
(void) free((char *)col0);
(void) free((char *)coll);
}

```



Table 1 (cont')

```

/*
 *
 * print() -- only routine visible outside this module
 */
static:
5  * getmat() -- trace back best path, count matches: print()
  * pr_align() -- print alignments of described in array pf[]: print()
  * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
  * numstr() -- put out a number line: dumpblock()
10 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
  * stars() -- put a line of stars: dumpblock()
  * stripname() -- strip any path and prefix from a sequence
  */

15 #include "nw.h"

#define SPC 3
#define P_LINE 256 /* maximum output line */
#define P_SPC 3 /* space between name or num and seq */

20 extern _day[26][26];
int olen; /* set output line length */
FILE *f; /* output file */

25 print() print
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx == fopen(outfile, "w")) == 0) {
30 fprintf(stderr, "%s: can't write %s\n", prog, outfile);
        cleanup(1);
    }
    fprintf(f, "%<first sequence: %s (length = %d)\n", names[0], len0);
    fprintf(f, "%<second sequence: %s (length = %d)\n", names[1], len1);
35 olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax <= len1 - 1) { /* leading gap in x */
40 pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
45 lx -= pp[1].spc;
    }
    if (dmax0 <= len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
50 }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
55 getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

Table 1 (cont')

```

/*
 * trace back the best path, count matches
 */
static
5 getmat(lx, ly, firstgap, lastgap)                                getmat
    int    lx, ly;
    int    firstgap, lastgap;
    /* "core" (minus endgaps) */
    /* leading trailing overlap */
{
    int    nm, i0, i1, siz0, siz1;
    char    outd[22];
    double    pct;
    register    n0, n1;
    register char    *p0, *p1;

    /* get total matches, score
    */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
    25         if (siz0) {
                p1++;
                n1++;
                siz0--;
            }
            else if (siz1) {
    30                 p0++;
                n0++;
                siz1--;
            }
            else {
    35                 if (xbm["p0-'A'"] & xbm["p1-'A'"])
                        nm++;
                if (n0++ == pp[0].n[i0])
                        siz0 = pp[0].n[i0]++;
    40                 if (n1++ == pp[1].n[i1])
                        siz1 = pp[1].n[i1]++;
                p0++;
                p1++;
            }
    45     }

    /* pct homology:
    * if penalizing endgaps, base is the shorter seq
    * else, knock off overhangs and take shorter core
    */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    55     pct = 100.0*(double)nm/(double)lx;
    fprintf(lx, "u");
    fprintf(lx, "%<d match%s in an overlap of %d: %.2f percent similarity'u",
        nm, (nm == 1)? "" : "es", lx, pct);
    60

```

**Table 1 (cont')**

```

fprintf(fx, "< gaps in first sequence: %d", gaps);
if (gaps) {
    (void) sprintf(outx, " (%d %s%s)",
        ngaps, (dna)? "base": "residue", (ngaps == 1)? "": "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gaps);
    if (gaps) {
        (void) sprintf(outx, " (%d %s%s)",
            ngaps, (dna)? "base": "residue", (ngaps == 1)? "": "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "in< score: %d (match == %d, mismatch == %d, gap penalty == %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "in< score: %d (Dayhoff PAM 250 matrix, gap penalty == %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base": "residue", (firstgap == 1)? "": "s",
            lastgap, (dna)? "base": "residue", (lastgap == 1)? "": "s");
    else
        fprintf(fx, "< endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];        /* jump index for a path */
static      nc[2];        /* number at start of current line */
static      ni[2];        /* current elem number -- for gapping */
static      siz[2];
static char *pe[2];        /* ptr to current element */
static char *po[2];        /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nm;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nm = strlen(name[names[i]));
        if (nm > lmax)
            lmax = nm;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = outf[i];
    }
}

```

pr\_align

Table 1 (cont')

```

for (nn = n1 = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
        * do we have more of this sequence?
        */
        if (!ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *pof[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *pof[i]++ = '-';
            siz[i]--;
        }
        else { /* we're putting a seq element
            */
            *pof[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);

            pof[i]++;
            ps[i]++;

            /*
            * are we at next gap for this seq?
            */
            if (n1[i] == pp[i].x[ji[i]]) {
                /*
                * we need to merge all gaps
                * at this location
                */
                siz[i] = pp[i].n[ji[i]]++;
                while (n1[i] == pp[i].x[ji[i]])
                    siz[i] += pp[i].n[ji[i]]++;
            }
            n1[i]++;
        }
    }
    if (++n1 == olen) { /* more && n1 */
        dumpblock();
        for (i = 0; i < 2; i++)
            pof[i] = out[i];
        n1 = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars; pr_align()
 */
static
dumpblock()
{
    register i;

    for (i = 0; i < 2; i++)
        *pof[i] = '\0';
}

```

dumpblock

Table 1 (cont')

...dumpblock

```

5      (void) puts("n", fx);
      for (i = 0; i < 2; i++) {
10         if ("out[i] && (*out[i] != ' ' || *(po[i]) != ' ') {
            if (i == 0)
                nums(i);
            if (i == 0 && *out[i])
                stars();
15         putchar(i);
            if (i == 0 && *out[i])
                fprintf(fx, star);
            if (i == 1)
                nums(i);
        }
    }

/*
20  * put out a number line: dumpblock()
  */
static
nums(ix)                                nums
25  {
    int      ix;      /* index in out[] holding seq line */

    char      nline[P_LINE];
    register  i, j;
    register  char  *pn, *px, *py;

30    for (pn = nline, i = 0; i < lmax + P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
35        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -1;
                for (px = pn; j; j /= 10, px--)
                    *px = j%10 + '0';
40                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
45                i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
50    for (pn = nline; *pn; pn++)
        (void) puts(*pn, fx);
    (void) puts("n", fx);
}

/*
55  * put out a line (name, [num], seq, [num]): dumpblock()
  */
static
putline(ix)                                putline
60  {
    int      ix;
    {

```

Table 1 (cont')

```

...putline
5      int          i;
      register char *px;

      for (px = namex[ix], i = 0; *px && *px != ' '; px++, i++)
          (void) puts(px, fx);
      for (; i < lmax+P_SPC; i++)
          (void) puts(' ', fx);

10     /* these count from 1:
       * n[] is current element (from 1)
       * w[] is number at start of current line
       */
15     for (px = out[ix], *px; px++)
          (void) puts(*px&0x7F, fx);
          (void) puts("\n", fx);
    }

20    /*
     * put a line of stars (segs always in out[0], out[1]); dumpblock()
     */
    static
    stars()
25    {
        int          i;
        register char *p0, *p1, cx, *px;

30        if (!out[0] || (!out[0] == ' ' && *out[0] == ' ') ||
            !out[1] || (!out[1] == ' ' && *out[1] == ' '))
            return;
        px = stat;
        for (i = lmax+P_SPC; i-->0)
35            *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
            if (strcmp(*p0-'A', *p1-'A') {
40                cx = '0';
                nm += 1;
            }
            else if (ldm && _day[*p0-'A'][*p1-'A'] > 0) {
45                cx = '.';
            }
            else
                cx = ' ';

            }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
55 }

```

Table 1 (cont')

```

/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
5  stripname(pn)                                stripname
{
    char    *pn;    /* file name (may be path) */

    register char    *px, *py;

10    py = 0;
    for (px = pn; px != '\0')
        if (*px == '/')
            py = px + 1;

15    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}

20

25

30

35

40

45

50

55

60

```

Table 1 (cont')

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * _g_calloc() -- calloc() with error checking
5  * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: mw()
 */
#include "aw.h"
#include <sys/file.h>

10 char    *jname = "/tmp/homgXXXXXX";    /* tmp file for jumps */
FILE     *fj;

int       cleanup();                    /* cleanup tmp file */
long      lseekf();

/*
 * remove any tmp file if we blow
 */
20 cleanup()
{
    int    i;

    if (fj)
        (void) unlink(jname);
25     exit(1);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ',', '<', or '>'
 * seq in upper or lower case
 */
30 char    *
getseq(file, len)                                getseq
35 char    *file;    /* file name */
    int    *len;    /* seq len */
{
    char    line[1024], *pseq;
    register char    *px, *py;
    int      nsize, then;
    FILE     *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
45     exit(1);
    }
    then = nsize = 0;
    while (!fgets(line, 1024, fp)) {
        if (*line == ',' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                then++;
    }
    if ((pseq = malloc((unsigned)(then+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, then+6, file);
55     exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
60

```



Table 1 (cont')

```

py = pseq + 4;
*len = ilen;
rewind(fp);

5   while (fgetc(line, 1024, fp)) {
        if (*line == ':' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
10          if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU",*(py-1)))
15              natgc++;
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (flen/3);
        return(jseq+4);
    }

25  char *
g_alloc(msg, nx, sz)                                g_alloc
    char *msg; /* program, calling routine */
    int nx, sz; /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
35        if (*msg) {
            fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
        return(px);
    }

40 }

/*
 * get final jumps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjumps()                                          readjumps
{
    int id = -1;
    int sz, id, il;
    register i, j, xx;

50     if (!id) {
        (void) fclose(f);
        if ((id = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
55         }
    }
    for (i = id == 0, dmax0 = dmax, xx = len0; i <+; i++) {
        while (1) {
60             for (j = dx[dmax].jmap; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;

```

**Table 1 (cont')**

...readjumps

```

5         if (j < 0 && dx[dmax].offset && fj) {
            (void) fseek(fd, dx[dmax].offset, 0);
            (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
            (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
            dx[dmax].ijmp = MAXJMP - 1;
        }
        else
            break;
10    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
15    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
20            pp[1].n[j] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1
               */
            pp[1].x[j] = xx - dmax + len1 - 1;
25            gapx ++;
            ngapx += siz;
        /* ignore MAXGAP when doing endgaps */
            siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
            if ++;
30        }
        else if (siz > 0) { /* gap in first seq */
            pp[0].n[j] = siz;
            pp[0].x[j] = xx;
            gapx ++;
            ngapx += siz;
35        /* ignore MAXGAP when doing endgaps */
            siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
            if ++;
40        }
    }
    else
        break;
}

/* reverse the order of jumps
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
50 }
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
55 }
if (fd >= 0)
    (void) close(fd);
if (!f) {
    (void) unlink(jname);
    fj = 0;
    offses = 0;
60 }
}

```

**Table 1 (cont')**

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
5  writejmps(ix)                                writejmps
    int    ix;
    {
        char    *mktemp();
10         if (!f) {
            if (mktemp(jname) < 0) {
                fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                cleanup(1);
            }
15         if ((f = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
20         (void) fwrite((char *)&dx[ix], sizeof(struct jmp), 1, f);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, f);
    }
30
35
40
45
50
55
60

```

**Table 2**

TAT	XXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXXXXXXX	(Length = 12 amino acids)

5      % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the TAT polypeptide) =

10      5 divided by 15 = 33.3%

**Table 3**

TAT	XXXXXXXXXX	(Length = 10 amino acids)
15      Comparison Protein	XXXXXXXXXXXXXX	(Length = 15 amino acids)

% amino acid sequence identity =

20      (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the TAT polypeptide) =

5 divided by 10 = 50%

**Table 4**

25

TAT-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

30

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the TAT-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

Table 5

TAT-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the TAT-DNA nucleic acid sequence) =

10 4 divided by 12 = 33.3%

## II. Compositions and Methods of the Invention

### A. Anti-TAT Antibodies

15 In one embodiment, the present invention provides anti-TAT antibodies which may find use herein as therapeutic and/or diagnostic agents. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

#### 1. Polyclonal Antibodies

20 Polyclonal antibodies are preferably raised in animals by multiple subcutaneous (sc) or intraperitoneal (ip) injections of the relevant antigen and an adjuvant. It may be useful to conjugate the relevant antigen (especially when synthetic peptides are used) to a protein that is immunogenic in the species to be immunized. For example, the antigen can be conjugated to keyhole limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, or soybean trypsin inhibitor, using a bifunctional or derivatizing agent, e.g., maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride,  $\text{SOCl}_2$ , or  $\text{R}^1\text{N}=\text{C}=\text{NR}$ , where R and  $\text{R}^1$  are different alkyl groups.

25 Animals are immunized against the antigen, immunogenic conjugates, or derivatives by combining, e.g., 100  $\mu\text{g}$  or 5  $\mu\text{g}$  of the protein or conjugate (for rabbits or mice, respectively) with 3 volumes of Freund's complete adjuvant and injecting the solution intradermally at multiple sites. One month later, the animals are boosted with 1/5 to 1/10 the original amount of peptide or conjugate in Freund's complete adjuvant by subcutaneous injection at multiple sites. Seven to 14 days later, the animals are bled and the serum is assayed for antibody titer. Animals are boosted until the titer plateaus. Conjugates also can be made in recombinant cell culture as protein fusions. Also, aggregating agents such as alum are suitably used to enhance the immune response.

#### 2. Monoclonal Antibodies

35 Monoclonal antibodies may be made using the hybridoma method first described by Kohler et al., *Nature*, 256:495 (1975), or may be made by recombinant DNA methods (U.S. Patent No. 4,816,567).

In the hybridoma method, a mouse or other appropriate host animal, such as a hamster, is immunized as described above to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the protein used for immunization. Alternatively, lymphocytes may be immunized *in vitro*. After immunization, lymphocytes are isolated and then fused with a myeloma cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)).

The hybridoma cells thus prepared are seeded and grown in a suitable culture medium which medium preferably contains one or more substances that inhibit the growth or survival of the unfused, parental myeloma cells (also referred to as fusion partner). For example, if the parental myeloma cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the selective culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine (HAT medium), which substances prevent the growth of HGPRT-deficient cells.

Preferred fusion partner myeloma cells are those that fuse efficiently, support stable high-level production of antibody by the selected antibody-producing cells, and are sensitive to a selective medium that selects against the unfused parental cells. Preferred myeloma cell lines are murine myeloma lines, such as those derived from MOPC-21 and MPC-11 mouse tumors available from the Salk Institute Cell Distribution Center, San Diego, California USA, and SP-2 and derivatives e.g., X63-Ag8-653 cells available from the American Type Culture Collection, Manassas, Virginia, USA. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozhor, J. Immunol., 133:3001 (1984); and Brodeur et al., Monoclonal Antibody Production Techniques and Applications, pp. 51-63 (Marcel Dekker, Inc., New York, 1987)).

Culture medium in which hybridoma cells are growing is assayed for production of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunosorbent assay (ELISA).

The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis described in Munson et al., Anal. Biochem., 107:220 (1980).

Once hybridoma cells that produce antibodies of the desired specificity, affinity, and/or activity are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)). Suitable culture media for this purpose include, for example, D-MEM or RPMI-1640 medium. In addition, the hybridoma cells may be grown *in vivo* as ascites tumors in an animal e.g., by i.p. injection of the cells into mice.

The monoclonal antibodies secreted by the subclones are suitably separated from the culture medium, ascites fluid, or serum by conventional antibody purification procedures such as, for example, affinity chromatography (e.g., using protein A or protein G-Sepharose) or ion-exchange chromatography, hydroxylapatite chromatography, gel electrophoresis, dialysis, etc.

DNA encoding the monoclonal antibodies is readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as *E. coli* cells, simian COS cells, Chinese Hamster Ovary (CHO) cells, or myeloma cells that do not otherwise produce antibody proteins, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. Review articles on recombinant expression in bacteria of DNA encoding the antibody include Skerra et al., Curr. Opinion in Immunol., 5:256-262 (1993) and Plückthun, Immunol. Revs., 130:151-188 (1992).

In a further embodiment, monoclonal antibodies or antibody fragments can be isolated from antibody phage libraries generated using the techniques described in McCafferty et al., Nature, 348:552-554 (1990). Clackson et al., Nature, 352:624-628 (1991) and Marks et al., J. Mol. Biol., 222:581-597 (1991) describe the isolation of murine and human antibodies, respectively, using phage libraries. Subsequent publications describe the production of high affinity (nM range) human antibodies by chain shuffling (Marks et al., BioTechnology, 10:779-783 (1992)), as well as combinatorial infection and *in vivo* recombination as a strategy for constructing very large phage libraries (Waterhouse et al., Nuc. Acids. Res., 21:2265-2266 (1993)). Thus, these techniques are viable alternatives to traditional monoclonal antibody hybridoma techniques for isolation of monoclonal antibodies.

The DNA that encodes the antibody may be modified to produce chimeric or fusion antibody polypeptides, for example, by substituting human heavy chain and light chain constant domain ( $C_H$  and  $C_L$ ) sequences for the homologous murine sequences (U.S. Patent No. 4,816,567; and Morrison, et al., Proc. Natl. Acad. Sci. USA, 81:6851 (1984)), or by fusing the immunoglobulin coding sequence with all or part of the coding sequence for a non-immunoglobulin polypeptide (heterologous polypeptide). The non-immunoglobulin polypeptide sequences can substitute for the constant domains of an antibody, or they are substituted for the variable domains of one antigen-combining site of an antibody to create a chimeric bivalent antibody comprising one antigen-combining site having specificity for an antigen and another antigen-combining site having specificity for a different antigen.

### 3. Human and Humanized Antibodies

The anti-TAT antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized

antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeven et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

The choice of human variable domains, both light and heavy, to be used in making the humanized antibodies is very important to reduce antigenicity and HAMA response (human anti-mouse antibody) when the antibody is intended for human therapeutic use. According to the so-called "best-fit" method, the sequence of the variable domain of a rodent antibody is screened against the entire library of known human variable domain sequences. The human V domain sequence which is closest to that of the rodent is identified and the human framework region (FR) within it accepted for the humanized antibody (Sims et al., *J. Immunol.*, 151:2296 (1993); Chodhia et al., *J. Mol. Biol.*, 196:901 (1987)). Another method uses a particular framework region derived from the consensus sequence of all human antibodies of a particular subgroup of light or heavy chains. The same framework may be used for several different humanized antibodies (Carter et al., *Proc. Natl. Acad. Sci. USA*, 89:4285 (1992); Presta et al., *J. Immunol.*, 151:2623 (1993)).

It is further important that antibodies be humanized with retention of high binding affinity for the antigen and other favorable biological properties. To achieve this goal, according to a preferred method, humanized antibodies are prepared by a process of analysis of the parental sequences and various conceptual humanized products using three-dimensional models of the parental and humanized sequences. Three-dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, i.e., the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen. In this way, FR residues can be selected and combined from the recipient and import sequences so that the desired antibody characteristic, such as increased affinity for the target antigen(s), is achieved. In general, the hypervariable



region residues are directly and most substantially involved in influencing antigen binding.

Various forms of a humanized anti-TAT antibody are contemplated. For example, the humanized antibody may be an antibody fragment, such as a Fab, which is optionally conjugated with one or more cytotoxic agent(s) in order to generate an immunoconjugate. Alternatively, the humanized antibody may be an intact antibody, such as an intact IgG1 antibody.

As an alternative to humanization, human antibodies can be generated. For example, it is now possible to produce transgenic animals (e.g., mice) that are capable, upon immunization, of producing a full repertoire of human antibodies in the absence of endogenous immunoglobulin production. For example, it has been described that the homozygous deletion of the antibody heavy-chain joining region ( $J_H$ ) gene in chimeric and germ-line mutant mice results in complete inhibition of endogenous antibody production. Transfer of the human germ-line immunoglobulin gene array into such germ-line mutant mice will result in the production of human antibodies upon antigen challenge. See, e.g., Jakobovits et al., *Proc. Natl. Acad. Sci. USA*, 90:2551 (1993); Jakobovits et al., *Nature*, 362:255-258 (1993); Bruggemann et al., *Year in Immunol.*, 7:33 (1993); U.S. Patent Nos. 5,545,806, 5,569,825, 5,591,669 (all of GenPharm); 5,545,807; and WO 97/17852.

Alternatively, phage display technology (McCafferty et al., *Nature* 348:552-553 [1990]) can be used to produce human antibodies and antibody fragments *in vitro*, from immunoglobulin variable (V) domain gene repertoires from nonimmunized donors. According to this technique, antibody V domain genes are cloned *in-frame* into either a major or minor coat protein gene of a filamentous bacteriophage, such as M13 or fd, and displayed as functional antibody fragments on the surface of the phage particle. Because the filamentous particle contains a single-stranded DNA copy of the phage genome, selections based on the functional properties of the antibody also result in selection of the gene encoding the antibody exhibiting those properties. Thus, the phage mimics some of the properties of the B-cell. Phage display can be performed in a variety of formats, reviewed in, e.g., Johnson, Kevin S. and Chiswell, David J., *Current Opinion in Structural Biology* 3:564-571 (1993). Several sources of V-gene segments can be used for phage display. Clackson et al., *Nature*, 352:624-628 (1991) isolated a diverse array of anti-oxazolone antibodies from a small random combinatorial library of V genes derived from the spleens of immunized mice. A repertoire of V genes from unimmunized human donors can be constructed and antibodies to a diverse array of antigens (including self-antigens) can be isolated essentially following the techniques described by Marks et al., *J. Mol. Biol.* 222:581-597 (1991), or Griffith et al., *EMBO J.* 12:725-734 (1993). See, also, U.S. Patent Nos. 5,565,332 and 5,573,905.

As discussed above, human antibodies may also be generated by *in vitro* activated B cells (see U.S. Patents 5,567,610 and 5,529,275).

#### 4. Antibody fragments

In certain circumstances there are advantages of using antibody fragments, rather than whole antibodies. The smaller size of the fragments allows for rapid clearance, and may lead to improved access to solid tumors.

Various techniques have been developed for the production of antibody fragments. Traditionally, these fragments were derived via proteolytic digestion of intact antibodies (see, e.g., Morimoto et al., *Journal*

of Biochemical and Biophysical Methods 24:107-117 (1992); and Brunnan et al., Science, 229:81 (1985)). However, these fragments can now be produced directly by recombinant host cells. Fab, Fv and ScFv antibody fragments can all be expressed in and secreted from *E. coli*, thus allowing the facile production of large amounts of these fragments. Antibody fragments can be isolated from the antibody phage libraries discussed above. Alternatively, Fab'-SH fragments can be directly recovered from *E. coli* and chemically coupled to form F(ab')<sub>2</sub> fragments (Carter et al., BioTechnology 10:163-167 (1992)). According to another approach, F(ab')<sub>2</sub> fragments can be isolated directly from recombinant host cell culture. Fab and F(ab')<sub>2</sub> fragment with increased in vivo half-life comprising a salvage receptor binding epitope residues are described in U.S. Patent No. 5,869,046. Other techniques for the production of antibody fragments will be apparent to the skilled practitioner. In other embodiments, the antibody of choice is a single chain Fv fragment (scFv). See WO 93/16185; U.S. Patent No. 5,571,894; and U.S. Patent No. 5,587,458. Fv and scFv are the only species with intact combining sites that are devoid of constant regions; thus, they are suitable for reduced nonspecific binding during in vivo use. scFv fusion proteins may be constructed to yield fusion of an effector protein at either the amino or the carboxy terminus of an scFv. See Antibody Engineering, ed. Borrebaeck, supra. The antibody fragment may also be a "linear antibody", e.g., as described in U.S. Patent 5,641,870 for example. Such linear antibody fragments may be monospecific or bispecific.

#### 5. Bispecific Antibodies

Bispecific antibodies are antibodies that have binding specificities for at least two different epitopes. Exemplary bispecific antibodies may bind to two different epitopes of a TAT protein as described herein. Other such antibodies may combine a TAT binding site with a binding site for another protein. Alternatively, an anti-TAT arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD3), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16), so as to focus and localize cellular defense mechanisms to the TAT-expressing cell. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express TAT. These antibodies possess a TAT-binding arm and an arm which binds the cytotoxic agent (e.g., saporin, anti-interferon-α, vinca alkaloid, ricin A chain, methotrexate or radioactive isotope hapten). Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g., F(ab')<sub>2</sub> bispecific antibodies).

WO 96/16673 describes a bispecific anti-ErbB2/anti-FcγRIII antibody and U.S. Patent No. 5,837,234 discloses a bispecific anti-ErbB2/anti-FcγRI antibody. A bispecific anti-ErbB2/Fcα antibody is shown in WO98/02463. U.S. Patent No. 5,821,337 teaches a bispecific anti-ErbB2/anti-CD3 antibody.

Methods for making bispecific antibodies are known in the art. Traditional production of full length bispecific antibodies is based on the co-expression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Milstein et al., Nature 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure. Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, and in Trautnecker et al., EMBO J. 10:3655-3659 (1991).

According to a different approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. Preferably, the fusion is with an Ig heavy chain constant domain, comprising at least part of the hinge, C<sub>H</sub>2, and C<sub>H</sub>3 regions. It is preferred to have the first heavy chain constant region (C<sub>H</sub>1) containing the site necessary for light chain bonding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host cell. This provides for greater flexibility in adjusting the mutual proportions of the three polypeptide fragments in embodiments when unequal ratios of the three polypeptide chains used in the construction provide the optimum yield of the desired bispecific antibody. It is, however, possible to insert the coding sequences for two or all three polypeptide chains into a single expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios have no significant effect on the yield of the desired chain combination.

In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology 121:210 (1986).

According to another approach described in U.S. Patent No. 5,731,168, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the C<sub>H</sub>3 domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g., tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g., alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies include cross-linked or "heteroconjugate" antibodies. For example, one of the antibodies in the heteroconjugate can be coupled to avidin, the other to biotin. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360, WO 92/200373, and EP 03089). Heteroconjugate antibodies may be made using any convenient cross-linking methods. Suitable cross-linking agents are well known in the art, and are disclosed in U.S. Patent No. 4,676,980, along with a number of cross-linking techniques.

Techniques for generating bispecific antibodies from antibody fragments have also been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent, sodium

arsenite, to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab' thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immunobilization of enzymes.

Recent progress has facilitated the direct recovery of Fab'-SH fragments from *E. coli*, which can be chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.*, 175: 217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.*, 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a V<sub>H</sub> connected to a V<sub>L</sub> by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (scFv) dimers has also been reported. See Gruber et al., *J. Immunol.*, 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutti et al., *J. Immunol.*, 147:60 (1991).

#### 6. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimide and those disclosed, for example, in U.S. Patent No. 4,676,980.

#### 7. Multivalent Antibodies

A multivalent antibody may be internalized (and/or catabolized) faster than a bivalent antibody by a cell expressing an antigen to which the antibodies bind. The antibodies of the present invention can be

multivalent antibodies (which are other than of the IgM class) with three or more antigen binding sites (e.g., tetravalent antibodies), which can be readily produced by recombinant expression of nucleic acid encoding the polypeptide chains of the antibody. The multivalent antibody can comprise a dimerization domain and three or more antigen binding sites. The preferred dimerization domain comprises (or consists of) an Fc region or a hinge region. In this scenario, the antibody will comprise an Fc region and three or more antigen binding sites amino-terminal to the Fc region. The preferred multivalent antibody herein comprises (or consists of) three to about eight, but preferably four, antigen binding sites. The multivalent antibody comprises at least one polypeptide chain (and preferably two polypeptide chains), wherein the polypeptide chain(s) comprise two or more variable domains. For instance, the polypeptide chain(s) may comprise  $VD1-(X1)_n-VD2-(X2)_n-Fc$ , wherein  $VD1$  is a first variable domain,  $VD2$  is a second variable domain,  $Fc$  is one polypeptide chain of an Fc region,  $X1$  and  $X2$  represent an amino acid or polypeptide, and  $n$  is 0 or 1. For instance, the polypeptide chain(s) may comprise: VH-CH1-flexible linker-VH-CH1-Fc region chain; or VH-CH1-VH-CH1-Fc region chain. The multivalent antibody herein preferably further comprises at least two (and preferably four) light chain variable domain polypeptides. The multivalent antibody herein may, for instance, comprise from about two to about eight light chain variable domain polypeptides. The light chain variable domain polypeptides contemplated here comprise a light chain variable domain and, optionally, further comprise a CL domain.

#### 8. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, e.g., so as to enhance antigen-dependent cell-mediated cytotoxicity (ADCC) and/or complement dependent cytotoxicity (CDC) of the antibody. This may be achieved by introducing one or more amino acid substitutions in an Fc region of the antibody. Alternatively or additionally, cysteine residue(s) may be introduced in the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp. Med., 176:1191-1195 (1992) and Shopes, B. J. Immunol., 148:2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff et al., Cancer Research 53:2560-2563 (1993). Alternatively, an antibody can be engineered which has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design 3:219-230 (1989). To increase the serum half life of the antibody, one may incorporate a salvage receptor binding epitope into the antibody (especially an antibody fragment) as described in U.S. Patent 5,739,277, for example. As used herein, the term "salvage receptor binding epitope" refers to an epitope of the Fc region of an IgG molecule (e.g., IgG<sub>1</sub>, IgG<sub>2</sub>, IgG<sub>3</sub>, or IgG<sub>4</sub>) that is responsible for increasing the *in vivo* serum half-life of the IgG molecule.

#### 9. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, a growth inhibitory agent, a toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunocjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Alewites fordii* proteins, diaphlin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, carcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomyein, and the tricothecenes. A variety of radiometalides are available for the production of radioconjugated antibodies. Examples include  $^{212}\text{Bi}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{90}\text{Y}$ , and  $^{188}\text{Re}$ . Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyl)propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCl), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis (p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminemepmaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

Conjugates of an antibody and one or more small molecule toxins, such as a calicheamicin, maytansinoids, a trichothene, and CC1065, and the derivatives of these toxins that have toxin activity, are also contemplated herein.

#### Maytansine and maytansinoids

In one preferred embodiment, an anti-TAT antibody (full length or fragments) of the invention is conjugated to one or more maytansinoid molecules.

Maytansinoids are mitototic inhibitors which act by inhibiting tubulin polymerization. Maytansine was first isolated from the east African shrub *Maytenus serrata* (U.S. Patent No. 3,896,111). Subsequently, it was discovered that certain microbes also produce maytansinoids, such as maytansinol and C-3 maytansinol esters (U.S. Patent No. 4,151,042). Synthetic maytansinol and derivatives and analogues thereof are disclosed, for example, in U.S. Patent Nos. 4,137,230; 4,248,870; 4,256,746; 4,260,608; 4,265,814; 4,294,757; 4,307,016; 4,308,268; 4,308,269; 4,309,428; 4,313,946; 4,315,929; 4,317,821; 4,322,348; 4,331,598; 4,361,650; 4,364,866; 4,424,219; 4,450,254; 4,362,663; and 4,371,533, the disclosures of which are hereby expressly incorporated by reference.

#### Maytansinoid-antibody conjugates

In an attempt to improve their therapeutic index, maytansine and maytansinoids have been conjugated to antibodies specifically binding to tumor cell antigens. Immunocjugates containing maytansinoids and their therapeutic use are disclosed, for example, in U.S. Patent Nos. 5,208,020, 5,416,064 and European Patent EP 0 425 235 B1, the disclosures of which are hereby expressly incorporated by reference. Liu *et al.*, Proc. Natl. Acad. Sci. USA 93:8618-8623 (1996) described immunocjugates comprising a maytansinoid designated DMI linked to the monoclonal antibody C242 directed against human colorectal cancer. The conjugate was

found to be highly cytotoxic towards cultured colon cancer cells, and showed antitumor activity in an *in vivo* tumor growth assay. Chari et al., Cancer Research 52:127-131 (1992) describe immunoconjugates in which a maytansinoid was conjugated via a disulfide linker to the murine antibody A7 binding to an antigen on human colon cancer cell lines, or to another murine monoclonal antibody TA.1 that binds the HER-2/*neu* oncogene. The cytotoxicity of the TA.1-maytansinoid conjugate was tested *in vitro* on the human breast cancer cell line SK-BR-3, which expresses  $3 \times 10^5$  HER-2 surface antigens per cell. The drug conjugate achieved a degree of cytotoxicity similar to the free maytansinoid drug, which could be increased by increasing the number of maytansinoid molecules per antibody molecule. The A7-maytansinoid conjugate showed low systemic cytotoxicity in mice.

#### Anti-TAT polypeptide antibody-maytansinoid conjugates (immunoconjugates)

Anti-TAT antibody-maytansinoid conjugates are prepared by chemically linking an anti-TAT antibody to a maytansinoid molecule without significantly diminishing the biological activity of either the antibody or the maytansinoid molecule. An average of 3-4 maytansinoid molecules conjugated per antibody molecule has shown efficacy in enhancing cytotoxicity of target cells without negatively affecting the function or solubility of the antibody, although even one molecule of toxin/antibody would be expected to enhance cytotoxicity over the use of naked antibody. Maytansinoids are well known in the art and can be synthesized by known techniques or isolated from natural sources. Suitable maytansinoids are disclosed, for example, in U.S. Patent No. 5,208,020 and in the other patents and nonpatent publications referred to hereinabove. Preferred maytansinoids are maytansinol and maytansinol analogues modified in the aromatic ring or at other positions of the maytansinol molecule, such as various maytansinol esters.

There are many linking groups known in the art for making antibody-maytansinoid conjugates, including, for example, those disclosed in U.S. Patent No. 5,208,020 or EP Patent 0 425 235 B1, and Chari et al., Cancer Research 52:127-131 (1992). The linking groups include disulfide groups, thioether groups, acid labile groups, photolabile groups, peptidase labile groups, or esterase labile groups, as disclosed in the above-identified patents, disulfide and thioether groups being preferred.

Conjugates of the antibody and maytansinoid may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). Particularly preferred coupling agents include N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP) (Carlisso et al., Biochem. J. 173:723-737 [1978]) and N-succinimidyl-4-(2-pyridyldithio)pentanoate (SPP) to provide for a disulfide linkage.

The linker may be attached to the maytansinoid molecule at various positions, depending on the type of the link. For example, an ester linkage may be formed by reaction with a hydroxyl group using conventional coupling techniques. The reaction may occur at the C-3 position having a hydroxyl group, the C-14 position modified with hydroxymethyl, the C-15 position modified with a hydroxyl group, and the C-20

position having a hydroxyl group. In a preferred embodiment, the linkage is formed at the C-5 position of myxanthinol or a myxanthinol analogue.

#### Calicheamicin

Another immunoconjugate of interest comprises an anti-TAT antibody conjugated to one or more calicheamicin molecules. The calicheamicin family of antibiotics are capable of producing double-stranded DNA breaks at sub-picomolar concentrations. For the preparation of conjugates of the calicheamicin family, see U.S. patents 5,712,374, 5,714,586, 5,739,116, 5,767,285, 5,770,701, 5,770,710, 5,773,001, 5,877,296 (all to American Cyanamid Company). Structural analogues of calicheamicin which may be used include, but are not limited to,  $\gamma_1^1$ ,  $\alpha_2^1$ ,  $\alpha_3^1$ , N-acetyl- $\gamma_1^1$ , PSAG and  $\theta_1^1$  (Himman et al., Cancer Research 53:3336-3342 (1993), Lode et al., Cancer Research 58:2925-2928 (1998) and the aforementioned U.S. patents to American Cyanamid). Another anti-tumor drug that the antibody can be conjugated is QFA which is an antifolate. Both calicheamicin and QFA have intracellular sites of action and do not readily cross the plasma membrane. Therefore, cellular uptake of these agents through antibody mediated internalization greatly enhances their cytotoxic effects.

#### Other cytotoxic agents

Other antitumor agents that can be conjugated to the anti-TAT antibodies of the invention include BCNU, streptozocin, vincristine and 5-fluorouracil, the family of agents known collectively LL-E33288 complex described in U.S. patents 5,053,394, 5,770,710, as well as esperamicins (U.S. patent 5,877,296).

Enzymatically active toxins and fragments thereof which can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, sarin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, diabin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, erotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin and the tricothecenes. See, for example, WO 93/21232 published October 28, 1993.

The present invention further contemplates an immunoconjugate formed between an antibody and a compound with nucleolytic activity (e.g., a ribonuclease or a DNA endonuclease such as a deoxyribonuclease; DNase).

For selective destruction of the tumor, the antibody may comprise a highly radioactive atom. A variety of radioactive isotopes are available for the production of radiolabeled anti-TAT antibodies. Examples include  $Al^{243}$ ,  $I^{131}$ ,  $I^{125}$ ,  $Y^{90}$ ,  $Re^{186}$ ,  $Re^{188}$ ,  $Sm^{153}$ ,  $Bi^{212}$ ,  $P^{32}$ ,  $Pb^{212}$  and radioactive isotopes of Lu. When the conjugate is used for diagnosis, it may comprise a radioactive atom for scintigraphic studies, for example  $^{90m}Tc$  or  $I^{123}$ , or a spin label for nuclear magnetic resonance (NMR) imaging (also known as magnetic resonance imaging, mri), such as iodine-123 again, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen-15, oxygen-17, gadolinium, manganese or iron.

The radio- or other labels may be incorporated in the conjugate in known ways. For example, the peptide may be biosynthesized or may be synthesized by chemical amino acid synthesis using suitable amino acid precursors involving, for example, fluorine-19 in place of hydrogen. Labels such as  $^{90m}Tc$  or  $I^{123}$ ,  $Re^{186}$ ,



Re<sup>186</sup> and In<sup>111</sup> can be attached via a cysteine residue in the peptide. Yttrium-90 can be attached via a lysine residue. The IODOGEN method (Fraker et al (1978) Biochem. Biophys. Res. Commun. 80: 49-57) can be used to incorporate iodine-123. "Monoclonal Antibodies in Immunoscintigraphy" (Chatal, CRC Press 1989) describes other methods in detail.

Conjugates of the antibody and cytotoxic agent may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipinidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Viretta et al., Science 238:1096 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026. The linker may be a "cleavable linker" facilitating release of the cytotoxic drug in the cell. For example, an acid-labile linker, peptidase-sensitive linker, photolabile linker, dimethyl linker or disulfide-containing linker (Chari et al., Cancer Research 52:127-131 (1992); U.S. Patent No. 5,208,020) may be used.

Alternatively, a fusion protein comprising the anti-TAT antibody and cytotoxic agent may be made, e.g., by recombinant techniques or peptide synthesis. The length of DNA may comprise respective regions encoding the two portions of the conjugate either adjacent one another or separated by a region encoding a linker peptide which does not destroy the desired properties of the conjugate.

In yet another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pre-targeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) which is conjugated to a cytotoxic agent (e.g., a radionucleotide).

#### 10. Immunoliposomes

The anti-TAT antibodies disclosed herein may also be formulated as immunoliposomes. A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030 (1980); U.S. Pat. Nos. 4,485,045 and 4,544,545; and WO97/38731 published October 23, 1997. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter.

Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin et al., *J. Biol. Chem.* 257:286-288 (1982) via a disulfide interchange reaction. A chemotherapeutic agent is optionally contained within the liposome. See Gabizon et al., *J. National Cancer Inst.* 81(19):1484 (1989).

#### B. TAT Binding Oligopeptides

5 TAT binding oligopeptides of the present invention are oligopeptides that bind, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. TAT binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7,  
 10 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard,  
 15 it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, e.g., U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., *Proc. Natl. Acad. Sci. U.S.A.*, 81:3998-4002 (1984); Geysen et al., *Proc. Natl. Acad. Sci. U.S.A.*, 82:178-182 (1985); Geysen et al., in *Synthetic Peptides as Antigens*, 130-149 (1986);  
 20 Geysen et al., *J. Immunol. Meth.*, 102:259-274 (1987); Schoofs et al., *J. Immunol.*, 140:611-616 (1988), Cwiria, S. E. et al. (1990) *Proc. Natl. Acad. Sci. USA*, 87:6378; Lowman, H.B. et al. (1991) *Biochemistry*, 30:10832; Clarkson, T. et al. (1991) *Nature*, 352: 624; Marks, J. D. et al. (1991), *J. Mol. Biol.*, 222:581; Kang, A.S. et al. (1991) *Proc. Natl. Acad. Sci. USA*, 88:8363, and Smith, G. P. (1991) *Current Opin. Biotechnol.*, 2:668).

25 In this regard, bacteriophage (phage) display is one well known technique which allows one to screen large oligopeptide libraries to identify member(s) of those libraries which are capable of specifically binding to a polypeptide target. Phage display is a technique by which variant polypeptides are displayed as fusion proteins to the coat protein on the surface of bacteriophage particles (Scott, J.K. and Smith, G. P. (1990) *Science* 249: 386). The utility of phage display lies in the fact that large libraries of selectively randomized  
 30 protein variants (or randomly cloned cDNAs) can be rapidly and efficiently sorted for those sequences that bind to a target molecule with high affinity. Display of peptide (Cwiria, S. E. et al. (1990) *Proc. Natl. Acad. Sci. USA*, 87:6378) or protein (Lowman, H.B. et al. (1991) *Biochemistry*, 30:10832; Clarkson, T. et al. (1991) *Nature*, 352: 624; Marks, J. D. et al. (1991), *J. Mol. Biol.*, 222:581; Kang, A.S. et al. (1991) *Proc. Natl. Acad. Sci. USA*, 88:8363) libraries on phage have been used for screening millions of polypeptides or  
 35 oligopeptides for ones with specific binding properties (Smith, G. P. (1991) *Current Opin. Biotechnol.*, 2:668). Sorting phage libraries of random mutants requires a strategy for constructing and propagating a large number of variants, a procedure for affinity purification using the target receptor, and a means of evaluating the results

of binding enrichments. U.S. Patent Nos. 5,223,409, 5,403,484, 5,571,689, and 5,663,143.

Although most phage display methods have used filamentous phage, lambdoid phage display systems (WO 95/34683; U.S. 5,627,024), T4 phage display systems (Ren, Z.-J. et al. (1998) Gene 215:439; Zhan, Z. (1997) CAN 33:534; Jiang, J. et al. (1997) can 128:44380; Ren, Z.-J. et al. (1997) CAN 127:215644; Ren, Z.-J. (1996) Protein Sci. 5:1833; Efimov, V. P. et al. (1995) Virus Genes 10:173) and T7 phage display systems (Smith, G. P. and Scott, J.K. (1993) Methods in Enzymology, 217, 228-257; U.S. 5,766,905) are also known.

Many other improvements and variations of the basic phage display concept have now been developed. These improvements enhance the ability of display systems to screen peptide libraries for binding to selected target molecules and to display functional proteins with the potential of screening these proteins for desired properties. Combinatorial reaction devices for phage display reactions have been developed (WO 98/14277) and phage display libraries have been used to analyze and control bimolecular interactions (WO 98/20169; WO 98/20159) and properties of constrained helical peptides (WO 98/20036). WO 97/35196 describes a method of isolating an affinity ligand in which a phage display library is contacted with one solution in which the ligand will bind to a target molecule and a second solution in which the affinity ligand will not bind to the target molecule, to selectively isolate binding ligands. WO 97/46251 describes a method of biopanning a random phage display library with an affinity purified antibody and then isolating binding phage, followed by a micropanning process using microplate wells to isolate high affinity binding phage. The use of *Staphylococcus aureus* protein A as an affinity tag has also been reported (Li et al. (1998) Mol Biotech., 9:187). WO 97/47314 describes the use of substrate subtraction libraries to distinguish enzyme specificities using a combinatorial library which may be a phage display library. A method for selecting enzymes suitable for use in detergents using phage display is described in WO 97/09446. Additional methods of selecting specific binding proteins are described in U.S. Patent Nos. 5,498,538, 5,432,018, and WO 98/15833.

Methods of generating peptide libraries and screening these libraries are also disclosed in U.S. Patent Nos. 5,723,286, 5,432,018, 5,580,717, 5,427,908, 5,498,530, 5,770,434, 5,734,018, 5,698,426, 5,763,192, and 5,723,323.

### C. TAT Binding Organic Molecules

TAT binding organic molecules are organic molecules other than oligopeptides or antibodies as defined herein that bind, preferably specifically, to a TAT polypeptide as described herein. TAT binding organic molecules may be identified and chemically synthesized using known methodology (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules are usually less than about 2000 daltons in size, alternatively less than about 1500, 750, 500, 250 or 200 daltons in size, wherein such organic molecules that are capable of binding, preferably specifically, to a TAT polypeptide as described herein may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening organic molecule libraries for molecules that are capable of binding to a polypeptide target are well known in the art (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules may be, for example, aldehydes, ketones, oximes, hydrazones, somcarbazones, carbazides, primary amines, secondary amines, tertiary amines, N-substituted hydrazines, hydrazides, alcohols,

ethers, thiols, thioethers, disulfides, carboxylic acids, esters, amides, ureas, carbamates, carbonates, ketals, thioketals, acetals, thioacetals, aryl halides, aryl sulfonates, alkyl halides, alkyl sulfonates, aromatic compounds, heterocyclic compounds, anilines, alkenes, alkynes, diols, amino alcohols, oxazolidines, oxazolidines, thiazolidines, thiazolidines, enamines, sulfonamides, epoxides, aziridines, isocyanates, sulfonyl chlorides, diazo compounds, acid chlorides, or the like.

5 D. Screening for Anti-TAT Antibodies, TAT Binding Oligopeptides and TAT Binding Organic Molecules With the Desired Properties

Techniques for generating antibodies, oligopeptides and organic molecules that bind to TAT polypeptides have been described above. One may further select antibodies, oligopeptides or other organic molecules with certain biological characteristics, as desired.

10 The growth inhibitory effects of an anti-TAT antibody, oligopeptide or other organic molecule of the invention may be assessed by methods known in the art, e.g., using cells which express a TAT polypeptide either endogenously or following transfection with the TAT gene. For example, appropriate tumor cell lines and TAT-transfected cells may be treated with an anti-TAT monoclonal antibody, oligopeptide or other organic molecule of the invention at various concentrations for a few days (e.g., 2-7) days and stained with crystal violet or MTT or analyzed by some other colorimetric assay. Another method of measuring proliferation would be by comparing  $^3\text{H}$ -thymidine uptake by the cells treated in the presence or absence of an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule of the invention. After treatment, the cells are harvested and the amount of radioactivity incorporated into the DNA quantitated in a scintillation counter. Appropriate positive controls include treatment of a selected cell line with a growth inhibitory antibody known to inhibit growth of that cell line. Growth inhibition of tumor cells *in vivo* can be determined in various ways known in the art. Preferably, the tumor cell is one that overexpresses a TAT polypeptide. Preferably, the anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule will inhibit cell proliferation of a TAT-expressing tumor cell *in vitro* or *in vivo* by about 25-100% compared to the untreated tumor cell, more preferably, by about 30-100%, and even more preferably by about 50-100% or 70-100%, in one embodiment, at an antibody concentration of about 0.5 to 30  $\mu\text{g}/\text{ml}$ . Growth inhibition can be measured at an antibody concentration of about 0.5 to 30  $\mu\text{g}/\text{ml}$  or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. The antibody is growth inhibitory *in vivo* if administration of the anti-TAT antibody at about 1  $\mu\text{g}/\text{kg}$  to about 100 mg/kg body weight results in reduction in tumor size or reduction of tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

30 To select for an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule which induces cell death, loss of membrane integrity as indicated by, e.g., propidium iodide (PI), trypan blue or 7AAD uptake may be assessed relative to control. A PI uptake assay can be performed in the absence of complement and immune effector cells. TAT polypeptide-expressing tumor cells are incubated with medium alone or medium containing the appropriate anti-TAT antibody (e.g., at about 10  $\mu\text{g}/\text{ml}$ ), TAT binding oligopeptide or TAT binding organic molecule. The cells are incubated for a 3 day time period. Following each treatment, cells are washed and aliquoted into 35 mm strainer-capped 12 x 75 tubes (1ml per tube, 3 tubes

per treatment group) for removal of cell clumps. Tubes then receive PI (10µg/ml). Samples may be analyzed using a FACSCAN® flow cytometer and FACSCONVERT® CellQuest software (Becton Dickinson). Those anti-TAT antibodies, TAT binding oligopeptides or TAT binding organic molecules that induce statistically significant levels of cell death as determined by PI uptake may be selected as cell death-inducing anti-TAT antibodies, TAT binding oligopeptides or TAT binding organic molecules.

To screen for antibodies, oligopeptides or other organic molecules which bind to an epitope on a TAT polypeptide bound by an antibody of interest, a routine cross-blocking assay such as that described in Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, Ed Harlow and David Lane (1988), can be performed. This assay can be used to determine if a test antibody, oligopeptide or other organic molecule binds the same site or epitope as a known anti-TAT antibody. Alternatively, or additionally, epitope mapping can be performed by methods known in the art. For example, the antibody sequence can be mutagenized such as by alanine scanning, to identify contact residues. The mutant antibody is initially tested for binding with polyclonal antibody to ensure proper folding. In a different method, peptides corresponding to different regions of a TAT polypeptide can be used in competition assays with the test antibodies or with a test antibody and an antibody with a characterized or known epitope.

#### B. Antibody-Dependent Enzyme Mediated Prodrug Therapy (ADEPT)

The antibodies of the present invention may also be used in ADEPT by conjugating the antibody to a prodrug-activating enzyme which converts a prodrug (e.g., a peptidyl chemotherapeutic agent, see WO81/01145) to an active anti-cancer drug. See, for example, WO 88/07378 and U.S. Patent No. 4,975,278.

The enzyme component of the immunoconjugate useful for ADEPT includes any enzyme capable of acting on a prodrug in such a way so as to convert it into its more active, cytotoxic form.

Enzymes that are useful in the method of this invention include, but are not limited to, alkaline phosphatase useful for converting phosphate-containing prodrugs into free drugs; arylsulfatase useful for converting sulfate-containing prodrugs into free drugs; cytosine deaminase useful for converting non-toxic 5-fluorocytosine into the anti-cancer drug, 5-fluorouracil; proteases, such as serrafin protease, thermolysin, subtilisin, carboxypeptidases and cathepsins (such as cathepsins B and L), that are useful for converting peptide-containing prodrugs into free drugs; D-alanylcarboxypeptidases, useful for converting prodrugs that contain D-amino acid substituents; carbohydrate-cleaving enzymes such as  $\beta$ -galactosidase and neuraminidase useful for converting glycosylated prodrugs into free drugs;  $\beta$ -lactamase useful for converting drugs derivatized with  $\beta$ -lactams into free drugs; and penicillin amidases, such as penicillin V amidase or penicillin G amidase, useful for converting drugs derivatized at their amine nitrogens with phenoxyacetyl or phenylacetyl groups, respectively, into free drugs. Alternatively, antibodies with enzymatic activity, also known in the art as "abzymes", can be used to convert the prodrugs of the invention into free active drugs (see, e.g., Massey, Nature 328:457-458 (1987)). Antibody-abzyme conjugates can be prepared as described herein for delivery of the abzyme to a tumor cell population.

The enzymes of this invention can be covalently bound to the anti-TAT antibodies by techniques well known in the art such as the use of the heterobifunctional crosslinking reagents discussed above. Alternatively, fusion proteins comprising at least the antigen binding region of an antibody of the invention linked to at least

a functionally active portion of an enzyme of the invention can be constructed using recombinant DNA techniques well known in the art (see, e.g., Neuberger et al., *Nature* 312:604-608 (1984).

F. Full-Length TAT Polypeptides

The present invention also provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as TAT polypeptides. In particular, cDNAs (partial and full-length) encoding various TAT polypeptides have been identified and isolated, as disclosed in further detail in the Examples below.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the TAT polypeptides and encoding nucleic acids described herein, in some cases, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

G. Anti-TAT Antibody and TAT Polypeptide Variants

In addition to the anti-TAT antibodies and full-length native sequence TAT polypeptides described herein, it is contemplated that anti-TAT antibody and TAT polypeptide variants can be prepared. Anti-TAT antibody and TAT polypeptide variants can be prepared by introducing appropriate nucleotide changes into the encoding DNA, and/or by synthesis of the desired antibody or polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the anti-TAT antibody or TAT polypeptide, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the anti-TAT antibodies and TAT polypeptides described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the antibody or polypeptide that results in a change in the amino acid sequence as compared with the native sequence antibody or polypeptide. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the anti-TAT antibody or TAT polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the anti-TAT antibody or TAT polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

Anti-TAT antibody and TAT polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with

a full length native antibody or protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the anti-TAT antibody or TAT polypeptide.

Anti-TAT antibody and TAT polypeptide fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating antibody or polypeptide fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired antibody or polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, anti-TAT antibody and TAT polypeptide fragments share at least one biological and/or immunological activity with the native anti-TAT antibody or TAT polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

	Original Residue	Exemplary Substitutions	Preferred Substitutions
20	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	gln
	Cys (C)	ser	ser
25	Gln (Q)	asn	asn
	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe;	
30		norleucine	leu
	Leu (L)	norleucine; ile; val;	
		met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
35	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
40	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe;	
		ala; norleucine	leu

Substantial modifications in function or immunological identity of the anti-TAT antibody or TAT polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining

(a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- (1) hydrophobic: norleucine, met, ala, val, leu, ile;
- (2) neutral hydrophilic: cys, ser, thr;
- (3) acidic: asp, glu;
- (4) basic: asn, gln, his, lys, arg;
- (5) residues that influence chain orientation: gly, pro; and
- (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., *Nucl. Acids Res.*, 13:4331 (1986); Zoller et al., *Nucl. Acids Res.*, 15:6487 (1987)], cassette mutagenesis [Wells et al., *Gene*, 34:315 (1985)], restriction selection mutagenesis [Wells et al., *Philos. Trans. R. Soc. London Ser. A*, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the anti-TAT antibody or TAT polypeptide variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, *Science*, 244:1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, *The Proteins*, (W.H. Freeman & Co., N.Y.); Chothia, *J. Mol. Biol.*, 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

Any cysteine residue not involved in maintaining the proper conformation of the anti-TAT antibody or TAT polypeptide also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking. Conversely, cysteine bond(s) may be added to the anti-TAT antibody or TAT polypeptide to improve its stability (particularly where the antibody is an antibody fragment such as an Fv fragment).

A particularly preferred type of substitutional variant involves substituting one or more hypervariable region residues of a parent antibody (e.g., a humanized or human antibody). Generally, the resulting variant(s) selected for further development will have improved biological properties relative to the parent antibody from which they are generated. A convenient way for generating such substitutional variants involves affinity maturation using phage display. Briefly, several hypervariable region sites (e.g., 6-7 sites) are mutated to generate all possible amino substitutions at each site. The antibody variants thus generated are displayed in



a monovalent fashion from filamentous phage particles as fusions to the gene III product of M13 packaged within each particle. The phage-displayed variants are then screened for their biological activity (e.g., binding affinity) as herein disclosed. In order to identify candidate hypervariable region sites for modification, alanine scanning mutagenesis can be performed to identify hypervariable region residues contributing significantly to antigen binding. Alternatively, or additionally, it may be beneficial to analyze a crystal structure of the antigen-antibody complex to identify contact points between the antibody and human TAT polypeptide. Such contact residues and neighboring residues are candidates for substitution according to the techniques elaborated herein. Once such variants are generated, the panel of variants is subjected to screening as described herein and antibodies with superior properties in one or more relevant assays may be selected for further development.

Nucleic acid molecules encoding amino acid sequence variants of the anti-TAT antibody are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of the anti-TAT antibody.

#### H. Modifications of Anti-TAT Antibodies and TAT Polypeptides

Covalent modifications of anti-TAT antibodies and TAT polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of an anti-TAT antibody or TAT polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of the anti-TAT antibody or TAT polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking anti-TAT antibody or TAT polypeptide to a water-insoluble support matrix or surface for use in the method for purifying anti-TAT antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazotacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propionimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the anti-TAT antibody or TAT polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the antibody or polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence anti-TAT antibody or TAT polypeptide (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence anti-TAT antibody or TAT polypeptide. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Glycosylation of antibodies and other polypeptides is typically either N-linked or O-linked. N-linked refers to the attachment of the carbohydrate moiety to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one of the sugars N-acetylgalactosamine, galactose, or xylose to a hydroxyl amino acid, most commonly serine or threonine, although 5-hydroxyproline or 5-hydroxylysine may also be used.

Addition of glycosylation sites to the anti-TAT antibody or TAT polypeptide is conveniently accomplished by altering the amino acid sequence such that it contains one or more of the above-described tripeptide sequences (for N-linked glycosylation sites). The alteration may also be made by the addition of, or substitution by, one or more serine or threonine residues to the sequence of the original anti-TAT antibody or TAT polypeptide (for O-linked glycosylation sites). The anti-TAT antibody or TAT polypeptide amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the anti-TAT antibody or TAT polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the anti-TAT antibody or TAT polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the anti-TAT antibody or TAT polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of anti-TAT antibody or TAT polypeptide comprises linking the antibody or polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337. The antibody or polypeptide also may be entrapped in microcapsules prepared, for example, by conservation techniques or by interfacial polymerization (for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively), in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nanoparticles and nanocapsules), or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Oslo, A., Ed., (1980).

The anti-TAT antibody or TAT polypeptide of the present invention may also be modified in a way to form chimeric molecules comprising an anti-TAT antibody or TAT polypeptide fused to another,

heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the anti-TAT antibody or TAT polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the anti-TAT antibody or TAT polypeptide. The presence of such epitope-tagged forms of the anti-TAT antibody or TAT polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the anti-TAT antibody or TAT polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 235:192-194 (1992)]; an  $\alpha$ -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Luiz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the anti-TAT antibody or TAT polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of an anti-TAT antibody or TAT polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH<sub>1</sub> and CH<sub>2</sub>, or the hinge, CH<sub>1</sub>, CH<sub>2</sub> and CH<sub>3</sub> regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

#### 1. Preparation of Anti-TAT Antibodies and TAT Polypeptides

The description below relates primarily to production of anti-TAT antibodies and TAT polypeptides by culturing cells transformed or transfected with a vector containing anti-TAT antibody- and TAT polypeptide-encoding nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare anti-TAT antibodies and TAT polypeptides. For instance, the appropriate amino acid sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the anti-TAT antibody or TAT polypeptide may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the desired anti-TAT antibody or TAT polypeptide.

## 1. Isolation of DNA Encoding Anti-TAT Antibody or TAT Polypeptide

DNA encoding anti-TAT antibody or TAT polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the anti-TAT antibody or TAT polypeptide mRNA and to express it at a detectable level. Accordingly, human anti-TAT antibody or TAT polypeptide DNA can be conveniently obtained from a cDNA library prepared from human tissue. The anti-TAT antibody- or TAT polypeptide-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding anti-TAT antibody or TAT polypeptide is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

Techniques for screening a cDNA library are well known in the art. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like <sup>32</sup>P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

## 2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for anti-TAT antibody or TAT polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example,  $\text{CaCl}_2$ ,  $\text{CaPO}_4$ , liposome-mediated and electroporation. Depending on the host

cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., *supra*, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., *Gene*, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, *Virology*, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., *J. Bact.*, 130:946 (1977) and Hsiao et al., *Proc. Natl. Acad. Sci. (USA)*, 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polymethine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., *Methods in Enzymology*, 185:527-537 (1990) and Mansour et al., *Nature*, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescens*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA prf3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA prf3 phoA E15 (argF-lac)169 degP ompT kan*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA prf3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

Full length antibody, antibody fragments, and antibody fusion proteins can be produced in bacteria, in particular when glycosylation and P<sub>0</sub> effector function are not needed, such as when the therapeutic antibody is conjugated to a cytotoxic agent (e.g., a toxin) and the immunocjugate by itself shows effectiveness in tumor cell destruction. Full length antibodies have greater half life in circulation. Production in *E. coli* is faster and more cost efficient. For expression of antibody fragments and polypeptides in bacteria, see, e.g., U.S. 5,648,237 (Carter et. al.), U.S. 5,789,199 (Joly et al.), and U.S. 5,840,523 (Simmons et al.) which

describes translation initiation region (TIR) and signal sequences for optimizing expression and secretion, these patents incorporated herein by reference. After expression, the antibody is isolated from the E. coli cell paste in a soluble fraction and can be purified through, e.g., a protein A or G column depending on the isotype. Final purification can be carried out similar to the process for purifying antibody expressed e.g., in CHO cells.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for anti-TAT antibody- or TAT polypeptide-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,539; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *Yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesei* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwannomyces* such as *Schwannomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylophilic yeasts are suitable herein and include, but are not limited to, yeasts capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylophilic, 269 (1982).

Suitable host cells for the expression of glycosylated anti-TAT antibody or TAT polypeptide are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells, such as cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* have been identified. A variety of viral strains for transfection are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure. Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen. Virol., 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/DHFR (CHO, Urlaub et al., Proc. Natl. Acad. Sci. USA 77:4216 (1980)); mouse sertoli cells (TM4, Maffier, Biol. Reprod., 23:243-251 (1980));

monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Maffei et al., Annals N.Y. Acad. Sci. 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors for anti-TAT antibody or TAT polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

### 3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding anti-TAT antibody or TAT polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The TAT may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the anti-TAT antibody- or TAT polypeptide-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, ponicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces*  $\alpha$ -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the  $2\mu$  plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical

nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the anti-TAT antibody- or TAT polypeptide-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the anti-TAT antibody- or TAT polypeptide-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the  $\beta$ -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding anti-TAT antibody or TAT polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4960 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

Anti-TAT antibody or TAT polypeptide transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the anti-TAT antibody or TAT polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein, and insulin).